UiO : Department of Mathematics
University of Oslo

Exchange Options in Financial Markets

Model Uncertainty and Pricing in Finite and Infinite Dimensional Spaces

Kenneth Ravn

Master's Thesis, Autumn 2019

This master's thesis is submitted under the master's programme *Mathematics*, with programme option *Mathematics for applications*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 60 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Submitted to University of Oslo, department of Mathematics, in partial fullfillment of the requirements for the degree of *Master of Science in Mathematics*.

Supervisor: Professor Fred Espen Benth, University of Oslo, department of mathematics, section for Stochastic analysis, finance, insurance and risk.

This thesis is typeset in Times New Roman, using MathTime Professional 2.

The font size is 12 pt, on the recommendation of the Department of Mathematics at the University of Oslo. The margin width is the LaTeXstandard margin width for 12 pt font sizes.

This page is intentionally left blank.

...denn ganz oben werden deine Sonnen
voll und glühend umgedreht.
Doch in dir ist schon begonnen,
was die Sonnen übersteht.[1]
R. M. Rilke

[1] ...a billion stars go spinning through the night,
blazing high above your head.
But in you is the presence that
will be, when all the stars are dead.

Translation by Stephen Mitchell

This page is intentionally left blank.

1 Abstract

Background: Modelling stochastic processes in infinite dimensions has proven itself to be a useful in the theory of interest rate and fixed income markets. This theory was extended to include power forward prices by Audet, Heiskanen, Keppo & Vehviläinen [3], and has since been extended by (among others) Benth & Krühner [12, 11].
Purpose: This thesis aims to extend the work done by (among others) Fred Espen Benth and Paul Krühner and others that has contributed to the various projects on electricity markets at the University of Oslo. We state and prove several results on option pricing in both finite and infinite spaces, and provide theory on model uncertainty.
Methods: Mostly results from classical stochastic analysis and functional analysis.
Results: Several results on the theory of pricing spread options in finite dimensions, a result on calendar spread options in infinite dimensions, results on put options and results on the put-call parity in infinite dimensions. We have also introduced the notion of convex risk measures to the Filipović space H_w and a new class of function spaces named the Filipović spaces, denoted F^p. Also, we have extended some results on the estimation of the eigenelements of the covariance operator in Hilbert spaces. Also, we have corrected a result made by Fred Espen Benth and Paul Krühner, which had an erroneous proof.
Conclusions: By a lot of elbow grease, tender loving care and a great collection of results from a vast array of papers, we have made a quite readable introduction to the theory of exchange options and model risk in financial markets.

2 Acknowledgements

Firstly, I am greatly indebted to Professor Fred Espen Benth (Department of mathematics, University of Oslo), for always having time for me, and having enough faith in me to propose such a difficult and open thesis. Professor Benth's constant good mood and vast mathematical knowledge has been of great help. I am also indebted to the brilliant mathematician and physicist Marius Jonsson (University of Cambridge), for proof-reading and for all the fruitful mathematical discussions over the past five-or-so years. I would also like to thank Dr. Kristina Dahl and Dr. Nacira Agram. The same goes for my colleagues in study hall 1001.

I would also like to thank all professors in the courses relevant to this thesis, including (but not limited to) Fred Espen Benth, Frank Proske, Tom Lindstrøm, Geir Ellingsrud and others.

Lastly, I would like to thank my brother, my father and his wife Hilde, for helping me rebuild my kitchen in January and February, making me not lose more time than I did in the start of this thesis.

Contents

This page is intentionally left blank.

3 Introduction

In financial markets, an *exchange option* is an option written on the difference of two underlying risky assets, whose values at time t may be denoted by $S_i(t)$, for $i = 1, 2$.[2] The holder of such an option will then have the right to be paid the difference $S_1(\tau) - S_2(\tau)$ at maturity time $\tau \geq 0$[3], provided $S_1(\tau) > S_2(\tau)$. Otherwise the option is worthless. The payoff of an exchange option is therefore $\max(0, S_1(t) - S_2(t))$. Exchange options are a special case of the class of options known as *spread options*, which are basically the same as exchange options, but treat the spread as a call (or put) option written on the spread with a strike price K. Spread options have payoff function $\max(0, S_1(t) - S_2(t) - K)$ when we look at them as a call option. The natural question to ask is then: what is the *fair* price of such options? In the case of exchange options where the underlying follows a geometric Brownian motion, this question was answered by Margrabe [53] in 1978. We will in this thesis give a brief review with a different proof before we prove similar results where the underlying is modelled differently. The results in the first chapter rely heavily on the seminal works by Black & Scholes [14] (including the extensions by Merton [57, 58, 59]), Black [13] and Margrabe [53].

Hilbert space valued diffusion processes (for which the Hilbert space valued Wiener processes are a special case of) were introduced to study the Dirichlet problem and to study parabolic equations for functions of infinitely many variables. The special case of infinite dimensional Ornstein-Uhlenbeck processes was introduced by Malliavin[4], but we refer to Da Prato & Zabczyk [65] and Applebaum [1] for introductions. Hilbert space-valued Wiener processes are quite similar to its finite dimensional cousins, they are more complex in the form that they rely on covariance operators rather than matrices.

It has been shown empirically, that it would be favorable to model the time dynamics of the forward curves in power markets as Hilbert space-valued processes, see e.g Benth & Krühner [11] and the references therein. This is partly due to a high degree of idiosyncratic risk in the power markets. It has been found a clear correlation structure between contracts with different delivery times. The Hilbert space chosen to do the modelling of the forward curve is the so-called Filipović space,

[2]one usually uses $S(\cdot)$ when the underlying is a stock, and $f(\cdot)$ when the underlying is a forward contract.

[3]The results in this thesis only considers European options, that is options which can only be exercised at the time of maturity. American options, options that can be continuously exercised until maturity will only be commented on briefly.

[4]The original article is allegedly called "stochastic calculus of variations and hypoelliptic operators", but I am unable to find it.

introduced by Filipović [34], which is essentially a separable Hilbert space of absolutely continuous functions. Let us now denote the forward price by $f(t, T)$, and assume that it has the following dynamics

$$\mathrm{d}f(t, T) = \alpha(t, T)\,\mathrm{d}t + \Sigma(t, T)\,\mathrm{d}W(t),$$

following the framework of Heath Jarrow and Morton. If we then introduce the Musiela parametrization $T = t + x$, we can derive the *Musiela stochastic differential equation*

$$\mathrm{d}f(t, T) = \left(\alpha(t, T) + \frac{\partial f(t, x)}{\partial x}\right)\,\mathrm{d}t + \Sigma(t, T)\,\mathrm{d}W(t).$$

In order to make sense of equations like this, we must work in an appropriate space containing the entire forward curve $\{f(t, x)|x, t \geq 0\}$. The Filipović space does exactly this.

In recent years the notion of model uncertainty has proven itself to be of utmost importance. It has been noted by Cont [25], that the New-York based derivatives unit of the Bank of Tokyo/Mitsubishi suffered a $83 million loss in 1997 because their internal pricing model overvalued a single portfolio of swaps and options on U.S interest rates. And a short time after NatWest Capital Markets lost £50 million due to a mispriced portfolio of U.K and German interest rate options, and that many of the reasons for these losses could be attributed to model risk. Model risk, and model uncertainty is especially relevant for spread options, since most of the options of this kind is not traded liquidly on an exchange, but is traded *over-the-counter*, which means that they need to priced (and hopefully priced correctly to avoid arbitrage) each time, and are not affected in the same way as more liquid instruments by supply and demand. We will review several approaches to measure model uncertainty, the most prominent being the notion of *convex risk measures*, for which a vast array of literature exist. We also review an extension of this idea to so-called *risk capturing functionals*, proposed by Bannör & Scherer [5].

3.1 About the proofs and my contributions

All proofs written in this thesis is my own. All results proven by someone else is referenced. To be completely clear, I have made a table of all the proven results in this thesis. The table contains a list of the results that are completely my own, both in statement and proofs, the results where the statement is made by someone else, but not proven[5], and the results where the statement and proof is made by

[5]Typically results like "...it then follows that" or "...which allows us to see that".

someone else, but I have edited and/or clarified and/or added *missing results*. The result where both the statement and proof are made by me is marked with (†). The table of results can be found in Appendix 14.4. In addition, section 12.2.1 is in its entirety my own. The same goes for the application of the results earlier results in section 12.2.2.

3.2 About references to other works and crossreferences

To avoid confusion, all references to results in other texts are written with **bold** letters, and cross-references will be written with normal letters. For example,

1. "We have from **Theorem 5.1.1** in Filipović [34] that the Filipović space H_w is a separable Hilbert space."

2. "We have for example seen in Theorem 9.13 how to price calendar spread options when the underlying takes values in a separable Hilbert space."

3.3 How this thesis is organized

We start with some preliminary results. I have included results that it seems like most other master's students in stochastic analysis did *not* know by heart (or at all). I have therefore, in an attempt to keep it as short as possible not included the definition of the stochastic integral, Itô's formula and related results. For these definitions and results we refer to Øksendal [62], Protter [66] and Benth [9].

After the preliminaries, we start with the problem of pricing spread options in $\mathbb{R}$ and $\mathbb{R}^n$. We start with the famous *Margrabe formula*, and then expand the idea of spread options to other situations. Some of these results are to the best of my knowledge new results. The theory on spread options in $\mathbb{R}$ is included as some form of *motivating* examples, motivating the pricing of spread options where the underlying follows an infinite dimensional representation. Since I was able to prove some new results, they are interesting as results on their own as well.

We then move on to some basic theory on Ornstein-Uhlenbeck processes (sometimes called the *Vasicek model*), and explain why such processes are the ones preferred when modelling spot prices in electricity (and related) markets. We do also include a brief section about infinite dimensional Ornstein-Uhlenbeck processes. Since infinite dimensional representations of Wiener processes is properly introduced later, I decided to keep this section brief, and use this section as a motivating example of an infinite dimensional stochastic process, since most of the master's students whose specialized in stochastic analysis know the class of

Ornstein-Uhlenbeck processes quite well.

The following section is on stochastic modelling of power markets. We introduce some of the most used models, and explain how we price contracts with a fixed delivery time, and contracts with a delivery period. The especially useful *Musiela parametrization* is also given a brief introduction is this section. The section on the Musiela parametrization, although brief, will also serve as a bridge between the earlier introduced theory on Ornstein-Uhlenbeck processes and the framework of Heath, Jarrow & Morton to the more generalized theory on forward curves introduced later.

Following this, we start with the infinite dimensional representations. We start by introducing the Filipović space, a separable Hilbert space introduced by Filipović in his PhD thesis, where he called it *the forward curve space*. This space will the Hilbert space most suitable for our purposes. We then, using the theory of (among others) Zabczyk, Da Prato, Peszat, Carmona and Tehranchi, introduce the theory of Hilbert space valued Wiener processes, and their associated covariance operators. We give a self contained introduction, where I have included some "easy-to-read" examples showing off these definitions. After showing off how stochastic equations in infinite dimensions may have several types of solutions depending on the initial conditions and *Lipschitzness* of the variables, we conclude this section with a brief discussion on how we can use the Hilbert space valued Wiener processes to price infinite dimensional representations of power forwards.

Then the main section of the thesis follows. In this section we expand on the end of the previous section, demonstrating and proving how one can price forward contracts with delivery over a time period, and the pricing of derivatives on these. Many of the results in this section are new, and one is a corrected result where the proof of the published version was wrong. We state and and prove results for both arithmetic and geometric models, and include a very brief discussions on how one would go about pricing such derivatives where the the noise of the underlying follows a Lévy process. This section follows closely a paper by Benth & Krühner[6] [11]. We show that in the same way as with the Black-Scholes formula (and related formulas), the price depends on the variance of some sort, and in this case the covariance operator.

Having shown that the price of the derivatives depends on the covariance operator,

[6]It is to be noted that Paul Krühner has since his joint works with Fred Espen Benth changed his name to Paul Eisenberg.

the following section is devoted to results on how one can specify and estimate the covariance operator, since the true covariance operator is not known in real-world situations. We start with a motivating discussion on parameter estimation on $\mathbb{R}$, where we discuss how to construct estimators for the quantities needed in Margrabe's formula. We also discuss their distributions. We then move on to the discussion on how to estimate the covariance operator. We include several results on this topic, including some new results on the estimation of the eigenelements of the covariance operator. This section follows Bosq's monograph **Linear Processes in Function Spaces** [15] closely.

In the following sections, we abandon the problem of pricing derivatives, and instead take a look on the uncertainty of the prices. The main tools of these sections are the so-called *risk measures*, mainly *convex risk measures*. In this section we include a discussion where we compare the models of several authors, including Cont's *worst case*-approach [25], and Bannör & Scherer's *risk capturing functionals*. We also include a brief discussion on how the notion of convex risk measures may be extended to other spaces, notably the Lebesgue spaces L^p consisting of p-integrable functions. We then extend the notions of convex risk measures to the Filipović space, which to the best of my knowledge has not been done before. Also, we construct an entire class of separable Banach spaces, which I have called the Filipović spaces, denoted F^p.

We end the thesis with a discussion and summary on how the prices of financial derivatives depend on the variance, covariance and/or correlation, and how model uncertainty impacts these prices.

Then a series of appendices follows. First an index of frequently used notation, then some basic results from mathematical finance we have referenced earlier, and sometimes could not find proofs for (like in the main part of the thesis, if a result is proven, it is proven by me). We then have some background materials on topological and normed spaces, and finally a list of the proofs and a list of the figures.

3.4 Why this thesis may be relevant for actuarial sciences

Being in possession of an adequate reserve is one of the main concerns for an insurance company. They must be able to meet their financial obligations, for example when paying for damages or other schemes where the insured has been guaranteed a certain amount of money now or in the future. One way to do this is just to offer an interest rate guarantee which is lower than the current interest rate in the market. This scheme is fine in times when the interest rates in the market are high. However, when interest rates fall beyond a certain threshold, this scheme becomes impossible, and the insurer has to speculate in the market in order to make enough money to pay the insurees. For example, on August 22, 2019, Finansavisen[7] wrote that the central bank of Denmark would lower the policy rate to -0.75, from the already negative current (on August 22) rate which was -0.65. Also, there are banks in other countries, for example Switzerland [8], that have negative interest rates on large deposits, making the aforementioned scheme even worse. This is where the theory on the financial markets may be of value, as we have proven several results on both regular call and put options (albeit in an infinite dimensional perspective) and results on spread and exchange options (both in finite dimensional and infinite dimensional settings), which may be used as an effective hedging tool in order to control future losses.

Delbaen, nne of the authors of the first paper defining risk measures in a financial-mathematical context, wrote in his (year) 2000 paper [27] the risk measure is denoted by ρ and a financial position by X, then $\rho(X)$ is the premium that needs to be charged in order to make position X insurable. This in turns means that if X is a position exhibiting extreme risk, i.e $\rho(A) = \infty$, then the position is unacceptable no matter how high the premium. In other words, "*These measures of risk can be used as (extra) capital requirements to regulate the risk assumed by market participants, traders, and insurance underwriters, as well as to allocate existing capital.*" - Artzner, Delbaen, Eber & Heath [2]. It is to be noted that if $\rho(X)$ is to interpreted as the margin requirement, we need that ρ is normalized in the sense that $\rho(0) = 0$.

Two commonly used risk measures are known as *Value-at-Risk*, abbreviated VaR, and *Average-Value-at-Risk*, abbreviated AVaR, although VaR does not satisfy the mathematical axioms in order to make it *convex* or *coherent* risk measure On the other hand, AVaR satisfy these. In this dissertation we will briefly introduce these two, but we refer to other works, for example Hull [44], for a more thorough introduction.

3.5 The work on this thesis

As a part of the master's curriculum, I took a course on stochastic modelling of electricity and related markets and one course on interest rates modelled as stochastic partial differential equations[9] taught by Professor Fred Espen Benth.

[7]Norwegian financial newspaper. The actual article can be found on page 18.
[8]Also found in Finansavisen, August 22, 2019, page 18.
[9]MAT4770 and STK4530, at UiO, department of mathematics.

Finding the content of these courses highly interesting (especially the first one), I asked Professor Benth about the supervising a master's thesis related to these courses. He replied with an idea where many of the ideas would be modelled in an infinite dimensional setting, and where I then could apply the ideas from the literature on model uncertainty. He then gave me the book "Stochastic equations in infinite dimensions" by Da Prato & Zabczyk [65], the papers by Benth & Krühner [12, 11], and told me to read it and to read up on the works by Rama Cont on model uncertainty. From there on, all the other literature has been found by me. This thesis is written independently, but has benefited greatly from Benth's helpful comments, especially when stuck. As with many others, I was shocked by the knowledge gap one has after completing the courses and starting reading research papers. However, it was good fun and the learning outcome has been huge on my own part.

All of the work on this thesis has been done between January 2019 and November 2019.

I am solely responsible for any errors.

Blindern, Oslo, the 14th of November 2019,
Kenneth Ravn.

4 Mathematical preliminaries

In his book **Spaces**, Lindstrøm [52] states *"Chapters with the word 'preliminaries' in the title are never much fun, but they are useful"*. This chapter is not an exception. I have tried to list *all* definitions and results that I feel should not be obvious for mathematicians (and in particular stochasticians) at my own level. This chapter is made up by two parts. One with results and definitions from real analysis and functional analysis, and one with results from stochastic analysis.

4.1 Results from real and functional analysis

Definition 4.1 (Banach Space)**.** A *Banach space* is a vector space X over a scalar field, which is equipped with a norm $\|\cdot\|_X$ and which is complete with respect to the metric induced by the norm. If X contains a countable dense subset it is called a separable Banach space.

Definition 4.2 (Hilbert Space)**.** An inner product space H is a called a Hilbert space if H is complete with respect to $||\cdot||_H$. If H admits a countable orthonormal basis it is called a separable Hilbert space.

Definition 4.3 (L^p-spaces)**.** Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, and let $1 \leq p < \infty$. The space $L^p(\Omega, \mathcal{A}, \mu)$ (often abbreviate $L^p(\mu)$ or simply L^p) consists of equivalence classes of measurable functions f such that

$$\int_\Omega |f|^p \, \mathrm{d}\mu < \infty,$$

equipped with the norm

$$\|f\|_p = \left(\int_\Omega |f|^p \, \mathrm{d}\mu\right)^{1/p}.$$

Theorem 4.4 (The Fubini-Tonelli theorem)**.** *Assume that S and T are σ-finite measure spaces and that f is a measurable function. Then*

$$\int_S \int_T |f(s,t)| \, \mathrm{d}t \, \mathrm{d}s = \int_T \int_S |f(s,t)| \, \mathrm{d}s \, \mathrm{d}t = \int_{S\times T} |f(s,t)| \, \mathrm{d}(s,t).$$

Moreover, if any one of the three integrals above is finite, then

$$\int_S \int_T f(s,t) \, \mathrm{d}t \, \mathrm{d}s = \int_T \int_S f(s,t) \, \mathrm{d}s \, \mathrm{d}t = \int_{S\times T} f(s,t) \, \mathrm{d}(s,t).$$

Definition 4.5 (Isometry)**.** Let X, Y be metric spaces equipped with metrics d_X and d_Y respectively. A map $f : X \to Y$ is said to an *isometry* if $d_Y(f(x), f(y)) = d_X(x, y)$. If f is bijective, it is called an *isometric isomorphism*.

Definition 4.6 (Closed linear operator)**.** Let X, Y be Banach spaces. A linear operator $A : D(A) \subset X \to Y$ is closed if for every sequence $\{x_n\}_n$ in $D(A)$ converging to $x \in X$ such that $Ax_n \to y \in Y$ we get $x \in D(A)$ and $Ax = y$.

Definition 4.7 (Hilbert-Schmidt operator)**.** A Hilbert Schmidt operator is a bounded operator A on a Hilbert space H with finite Hilbert-Schmidt norm:

$$||A||_{\text{HS}}^2 = \text{Tr}\,(A^* A) = \sum_{i \in J} ||Ae_i||^2,$$

where J is some (possibly uncountable) index set, $||\cdot||$ is the norm of H and $\{e_i\}_{i \in I}$ is an orthonormal basis of H.

Definition 4.8 (Semigroup)**.** Let X be a Banach space. A family $(T(t), t \geq 0)$ of bounded linear operators on X is a called a strongly continuous semigroup if

1. $T(0) = I$, where I denotes the identity operator on X
2. $T(t+s) = T(t)T(s)$, $\forall\, t, s \geq 0$
3. For all $x_0 \in X$, $||T(t)x_0 - x_0|| \to 0$ whenever $t \downarrow 0$.

Definition 4.9 (Generator of semigroup)**.** Let X be a Banach space. The generator $A : D(A) \subset X \to X$ of a strongly continuous semigroup on X is the operator

$$Ax = \lim_{h \downarrow 0} \frac{1}{h}\left(T(h)x - x\right). \tag{1}$$

(1) is defined for all x in the domain of A:

$$D(A) = \left\{x \in X : \lim_{h \downarrow 0} \frac{1}{h}\left(T(h)x - x\right) \text{ exists}\right\}. \tag{2}$$

Definition 4.10 (Covariance operator)**.** Assume X is a square integrable random variable defined on H. Then $\mathcal{Q} \in L(H)$ is called the covariance operator of X if

$$\mathbb{E}\left[\langle X, u\rangle\langle X, v\rangle\right] = \langle \mathcal{Q}u, v\rangle.$$

Theorem 4.11 (Cauchy-Schwarz' inequality/Cauchy-Schwarz-Bunyakovsky's inequality)**.** *For any vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in an inner product space, it holds true that*

$$|\langle \mathbf{u}, \mathbf{v}\rangle|^2 \leq \langle \mathbf{u}, \mathbf{u}\rangle \cdot \langle \mathbf{v}, \mathbf{v}\rangle. \tag{3}$$

Theorem 4.12 (Jordan decomposition)**.** *Let* $(\Omega, \mathcal{A})$ *be a measurable space and* ν *a signed measure on* $\mathcal{A}$*. Then* ν *can be expressed uniquely as* $\nu = \nu_+ - \nu^-$*, where* ν^+ *and* ν^- *are two mutually singular measures on* $\mathcal{A}$*. Two measures are mutually singular if there exists some* $E \in \mathcal{A}$ *such that* $\nu^+(E^c) = 0$ *and* $\nu^-(E) = 0$*.*

Definition 4.13 (Absolutely continuous measure)**.** Two measures on a measurable space $(\Omega, \mathcal{A})$ are *absolutely continuous* denoted $\mu \ll \nu$ if $\mu(A) = 0$ whenever $\nu(A) = 0$.

Theorem 4.14 (Radon-Nikodym theorem)**.** *Let $(\Omega, \mathcal{A}, \mu)$ be a σ-finite measure space and ν a σ-finite measure on $\mathcal{A}$. If $\nu \ll \mu$ then there is a non-negative extended real valued function f on Ω such that*

$$\nu(A) = \int_A f \, \mathrm{d}\mu, \ \ A \in \mathcal{A}.$$

Moreover, f is unique in the sense that if g has the same properties as f and $\nu(A) = \int_A g \, \mathrm{d}\mu$, then $f = g$ μ-almost everywhere. The quantity $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} = f$ is called the Radon-Nikodym derivative.

Theorem 4.15 (Urysohn's Metrization Theorem)**.** *A topological space is separable and metrizable if and only if it is regular, Hausdorff and second-countable.*

Definition 4.16 (Bochner integral)**.** Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, and B a Banach space. A measurable function $f : \Omega \to B$ is *Bochner integrable* if there exists a sequence of integrable simple functions $\{s_n\}_{n\geq 0}$ such that

$$\lim_{n\to\infty} \int_\Omega \|f - s_n\|_B \, \mathrm{d}\mu = 0,$$

where the integral on the left hand side is an ordinary Lebesgue integral (see e.g McDonald & Weiss [55]).

Then, the *Bochner integral* is defined by

$$\int_\Omega f \, \mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega s_n \, \mathrm{d}\mu.$$

We note that when it comes to Bochner integrals, the Radon-Nikodym theorem does not in general hold. However, due to the Dunford-Pettis theorem (Dellacherie & Meyer [28]) it holds in separable dual spaces, and also in Hilbert spaces (in fact all reflexive spaces, that is, all locally convex topological vector spaces that coincides with the continuous dual of its continuous dual space).
The next result is an extension of the theorem above of conditional expectations, and is due to Da Prato & Zabczyk [65].

Definition 4.17 (Measures of weak and strong order)**.** Let X be a Banach space and $0 < p < \infty$. A measure μ on $\hat{C}(X)$, where $\hat{C}(X)$ denotes the cylindrical σ-algebra on X, is said to be of *weak order p* if $\int |\langle x, x^* \rangle|^p \mu(\mathrm{d}x) < \infty$ and of *strong order p* if $\int \|x\|^p \mu(\mathrm{d}x) < \infty$.

Theorem 4.18 (Spectral theorem for self-adjoint compact operators)**.** *For every compact self-adjoint operator T on a real or complex Hilbert space H, there exists an orthonormal basis of H consisting of eigenvectors of T. More specifically, the orthogonal complement of the kernel of T admits, either a finite orthonormal basis of eigenvectors of T, or a countably infinite orthonormal basis $\{e_n\}$ of eigenvectors of T, with corresponding eigenvalues $\{\lambda_n\} \subset \mathbb{R}$, such that $\lambda_n \to 0$.*

Moreover, when H is separable and infinite dimensional we have that for every compact self-adjoint operator T on a real or complex separable infinite- dimensional Hilbert space H, there exists a countably infinite orthonormal basis $\{v_n\}$ of H consisting of eigenvectors of T, with corresponding eigenvalues $\{\lambda_n\} \subset \mathbb{R}$, such that $\lambda_n \to 0$.

Theorem 4.19 (Riesz-Fréchet representation theorem)**.** *Let H be a Hilbert space and $f \in H^*$. Then there exists a $g \in H$ such that $f(x) = \langle g, x \rangle$ for each $x \in H$. Moreover, $\|g\|_H = \|f\|_{H^*}$.*

4.2 Results from stochastic analysis

Definition 4.20 (Adaptedness)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_t)_{t\in I}$ be a filtration of the σ-algebra $\mathcal{F}$, where I is some index set, typically $[0, T]$ in our case. Furthermore, let (S, Σ) be a measurable space. A stochastic process $X : I \times \Omega \to S$ is said to be *adapted to the filtration $\mathcal{F}_t)_{t\in I}$* if the random variable $X_t : \Omega \to S$ is $(\mathcal{F}_t,)$-measurable for each $t \in I$.

Definition 4.21 (Levy Processes)**.** Let $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, P)$ be a filtered probability space, and let H be a separable Hilbert space. An adapted H-valued process $(L(t), t \geq 0))$ is called a Levy process if

1. $L(0) = 0$ P-a.s,
2. L is stochastically continuous,
3. L has càdlàg paths, that is right continuous with left limits
4. L has stationary increments, i.e for $t \geq s$ $L(t) - L(s) = L(t - s)$ in distribution,
5. L has independent increments, i.e for $t \geq s$ $L(t) - L(s)$ is independent of $\mathcal{F}_s$.

The next result is a stochastic version of Fubini's theorem. We refer to Filipović [33] for a more general statement and proof.

Theorem 4.22 (stochastic Fubini)**.** *Consider the* $\mathbb{R}^d$*-valued stochastic process* $\phi = \phi(\omega, t, s)$ *with two indexes* $0, \leq t$ *and* $s \leq T$ *satisfying the following properties*

(a) ϕ *is* $\mathcal{F}_T \otimes \mathcal{B}[0, T]$*-measurable*

(b) $\sup_{t,s} \|\phi(t, s)\| < \infty$.

Then

$$\int_0^T \int_0^T \phi(t, s)\, \mathrm{d}W(t)\, \mathrm{d}s = \int_0^T \int_0^T \phi(t, s)\, \mathrm{d}s\, \mathrm{d}W(t). \tag{4}$$

Definition 4.23 (Predictable processes)**.** Given a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$, a stochastic process $(X(t))_{t\geq 0}$ is *predictable* if X, considered as a mapping from $\Omega \times \mathbb{R}_+$ is measurable with respect to the σ-algebra generated by all left-continuous adapted processes.

The next theorem is one of the more famous theorems from stochastic analysis: Girsanov's theorem. We use a slightly modified version of the statement from Øksendal [62], and is therefore stated here.

Theorem 4.24 (Girsanov's Theorem)**.** *Let* $Y(t)$ *be an Itô process of the form*

$$\mathrm{d}Y(t) = -a(t, \omega)\, \mathrm{d}t + \mathrm{d}B(t), \quad t \leq T, Y_0 = 0,$$

where $T \leq \infty$ *is a given constant and* $B(t)$ *is a* n*-dimensional Brownian motion. Define*

$$M(t) = \exp\left(a(t, \omega) B(t)\right),$$

and the measure Q *on* $\mathcal{F}_T$ *by*

$$\frac{\mathrm{d}Q(\omega)}{\mathrm{d}P(\omega)} = M(T).$$

Then Q *is a probability measure on* $\mathcal{F}_T$ *and* $Y(t)$ *is an* n*-dimensional Brownian motion with respect to* Q.

Definition 4.25 (Expected value)**.** Let X be a random variable whose cumulative distribution function admits a density $f(x)$. The expected value of X, denoted $\mathbb{E}[X]$ is then defined as

$$\int_\Omega x f(x)\, \mathrm{d}x.$$

In general, if X is defined on a probability space $(\Omega, \mathcal{F}, P)$, then

$$\mathbb{E}[X] = \int_\Omega X(\omega)\, \mathrm{d}P(\omega).$$

Theorem 4.26 (Conditional expectations)**.** *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $Y \in L^1(\Omega, \mathcal{A}, P)$ and $\mathcal{F}$ a σ-algebra such that $\mathcal{F} \subset \mathcal{A}$. Then there exists a P-almost everywhere unique $\mathcal{F}$-measurable function called the conditional expectation of Y given $\mathcal{F}$ $\mathbb{E}[Y|\mathcal{F}]$ such that*

$$\int_F Y \, \mathrm{d}P = \int_F \mathbb{E}[Y|\mathcal{F}] \, \mathrm{d}P,$$

for all $F \in \mathcal{F}$.

Moreover, if $\mathbb{E}[|Y|] < \infty$, then with probability one

1. $\mathbb{E}[aX + bY|\mathcal{F}] = a\,\mathbb{E}[X|\mathcal{F}] + b\,\mathbb{E}[Y|\mathcal{F}]$
2. $E[E[Y|\mathcal{F}]] = E[Y]$
3. $\mathbb{E}[Y|\mathcal{F}] = Y$ *if Y is $\mathcal{F}$-measurable*
4. $\mathbb{E}[X \cdot Y|\mathcal{F}] = X\,\mathbb{E}[Y|\mathcal{F}]$ *if X is $\mathcal{F}$-measurable, where $\cdot$ denotes the usual inner product on $\mathbb{R}^n$*
5. $\mathbb{E}[Y|\mathcal{F}] = \mathbb{E}[Y]$ *if Y is independent of $\mathcal{F}$.*

We refer to McDonald & Weiss [55] and Øksendal [62] for more on conditional expectations.

Theorem 4.27 (Conditional expectations on Banach spaces)**.** *Assume E is a separable Banach space. Let X be a Bochner integrable E-valued random variable defined on $(\Omega, \mathcal{F}, P)$ and let $\mathcal{G} \subset \mathcal{F}$ be a sub σ-algebra. Then there exists a unique, up to a set of P-probability zero, integrable E-valued random variable Z, measurable with respect to $\mathcal{G}$ such that for all $G \in \mathcal{G}$*

$$\int_G X \, \mathrm{d}P = \int_G Z \, \mathrm{d}P.$$

The random variable Z will be denoted as $\mathbb{E}[X|\mathcal{G}]$ and is called the conditional expectation of X given $\mathcal{G}$.

Definition 4.28 (Bayes' formula for conditional expectations)**.** Let s and t satisfying $0 \le s \le t \le T$ be given, and let Y be an $\mathcal{F}_t$-measurable random variable. Further, define the Radon-Nikodym derivative

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_t} = M(t).$$

Then

$$\mathbb{E}_Q[Y|\mathcal{F}_s]\,\mathbb{E}[M(t)|\mathcal{F}_s] = \mathbb{E}[YM(t)|\mathcal{F}_s]. \tag{5}$$

5 Pricing spread options on the real line

Spread trading is a well known trading strategy. This strategy involves the purchase of one security while simultaneously selling another. The trades are executed in a certain manner to yield an overall net position on the trade, where the profit (or loss) comes from the widening or narrowing of the spread. In order to hedge oneself against future risks on spread trades, buying (or selling) options *written on* the spreads may be a feasible strategy. In this section we will derive and prove results on the pricing of such options. We will consider *exchange options*, whose valuation was proven by Margrabe in his seminal 1978 paper [53], both in the context of Margrabe's paper (but with another derivation) and in the case where we assume that the dynamics of the spread follows an arithmetic Brownian motion. The latter is called *Bachelier's model*, and its price was found by Carmona & Durrleman [17]. In this thesis we have extended their proof. We will then move on to *calendar spread options*, which are options written on the spread of two forward contracts on the same commodity, but with different delivering times. This is done in the setting of Heath, Jarrow & Morton [42] and in a setting where the underlying assets follows an n-dimensional arithmetic Brownian motion. We end this section with a brief review of an approximation formula for spread options, proved by Carmona & Durrleman [17].

We begin by defining the stochastic exponential, also known as the Doléans-Dade exponential. We use the definition from Protter [66].

Definition 5.1 (Doléans-Dade exponential)**.** For a semimartingale X with $X(0) = 0$, the *stochastic exponential* of X, denoted by $\mathcal{E}(X)$, is the unique semimartingale Z that is a solution of $Z(t) = 1 + \int_0^t Z(s)\,\mathrm{d}X(s)$.

Proposition 5.2. *If X is a continuous semimartingale, then*

$$\mathcal{E}(X)(t) = \exp\left(X(t) - \frac{1}{2}[X, X](t)\right).$$

Proof: See Protter [66], page 85. ∎

Lemma 5.3. *Let $\{W(t, t \geq 0)\}$ be a Brownian motion, and let $\sigma \in \mathbb{R}$. Then*

1. $\mathbb{E}[\mathcal{E}(\sigma W)(t)] = 1$
2. $\mathrm{Var}[\mathcal{E}(\sigma W)(t)] = e^{\sigma^2 t} - 1.$

Proof: Using the fact that a Brownian motion is a Gaussian process, whose exponential moments are finite, we get

$$\begin{aligned}
\mathbb{E}[\mathcal{E}(\sigma W)(t)] &= \exp\left(\mathbb{E}[W(t) - \frac{1}{2}\sigma^2 t] + \frac{1}{2}\mathrm{Var}[\sigma W(t)]\right) \\
&= \exp(-\frac{1}{2}\sigma^2 t + \frac{1}{2}\sigma^2 t) \\
&= 1.
\end{aligned}$$

In a similar fashion we find that

$$\begin{aligned}
\mathrm{Var}[\mathcal{E}(\sigma W)(t)] &= \mathrm{Var}\left[e^{\sigma W(t) - \frac{1}{2}\sigma^2 t}\right] \\
&= \mathbb{E}[(\mathcal{E}(W)(t))^2] - \mathbb{E}[\mathcal{E}(W)(t)]^2 \\
&= \mathbb{E}\left[e^{2\sigma W(t) - \sigma^2 t}\right] - 1^2 \\
&= e^{\sigma^2 t} - 1.
\end{aligned}$$

■

5.1 The option to exchange one asset for another - Margrabe's formula and some variations

In this section we will state and prove Margrabe's formula [53] for spread options where the strike price is zero. We provide a proof different to the one in Margrabe's paper.

We start with a technical result.

Lemma 5.4 (†). *Let Q be a probability measure from Girsanov with associated Brownian motion $\{W(t)\}_t$, generated from the P-Brownian motion $B(t)$. Moreover, let $Y(t)$ be a P-Brownian motion independent of $B(t)$. Then $Y(t)$ remains a Brownian motion under Q.*

Proof: This result follows from Lévy's Characterization of Brownian motion. We find that for an integrable stochastic variable X that $\mathbb{E}_Q[X] = \mathbb{E}[XM(t)]$, where $M(\cdot)$ is the Radon-Nikodym derivative that defines the measure Q in Girsanov's theorem. Then, using Bayes' formula for conditional expectations (5) and the fact that in Girsanov's theorem $t \mapsto M(t)$ is a P-martingale, we get that

$$\mathbb{E}_Q[Y(t)|\mathcal{F}_s] = \frac{\mathbb{E}[Y(t)M(t)|\mathcal{F}_s]}{\mathbb{E}[M(t)|\mathcal{F}_s]} \tag{6}$$

We then know that $M(t)$ is a stochastic exponential on the form $M(t) = \mathcal{E}(\eta B(t))$, for some η. And since we have that Y was independent of B by hypothesis, get that

$$\frac{\mathbb{E}[Y(t)M(t)|\mathcal{F}_s]}{\mathbb{E}[M(t)|\mathcal{F}_s]} = \frac{\mathbb{E}[Y(t)|\mathcal{F}_s]\,\mathbb{E}[M(t)|\mathcal{F}_s]}{\mathbb{E}[M(t)|\mathcal{F}_s]} = Y(s) \tag{7}$$

The first equality follows from the fact that we assumed that Y was independent of B, and the second equality follows from the fact that a Brownian motion is a martingale. From this we get that $t \mapsto Y(t)$ is a martingale under the measure Q. Moreover, by the same argument, $Y(t)^2 - t$ is a martingale under Q as well, and therefore we get that $Y(t)$ is a Brownian motion under Q by Lévy's characterization of Brownian motion. ■

We can now state and prove Margrabe's formula. This seminal result was proved by William Margrabe in 1978. We will give a different proof.

Proposition 5.5. *Consider two commodities, $S^{(1)}(t)$ and $S^{(2)}(t)$, with price dynamics given by a geometric Brownian motion. We will for simplicity assume that the drift term is constantly zero, and that the interest rate is zero as well. The price of a spread option as time $t = 0$ is then*

$$V(0) = S^{(1)}(0)\Phi(d_1) - S^{(2)}(0)\Phi(d_2), \tag{8}$$

where

$$d_1 = \frac{\log\left(S^{(1)}(0)/S^{(2)}(0)\right) + \frac{1}{2}t\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2\rho\right)}{\sqrt{t\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma^2\rho\right)}}$$

$$d_2 = d_1 - \sqrt{t\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma^2\rho\right)}.$$

The quantities $\sigma_1, \sigma_2 \geq 0$, $\rho \in [-1, 1]$ and Φ is the standard normal cumulative distribution function.

Proof: We then get:

$$\mathrm{d}S^{(i)}(t) = \sigma_i S^{(i)}(t)\,\mathrm{d}B^{(i)}(t), \tag{9}$$

hence

$$S^{(i)}(t) = S^{(i)}(0)\mathcal{E}\left(\sigma_i B^{(i)}(t)\right). \tag{10}$$

Here, $\left(B^{(i)}(t), t \geq 0, i = 1, 2\right)$ are two Brownian motions with correlation ρ.

We are interested in the value of an option to exchange one asset for another, that is we want to compute

$$V(0) = \mathbb{E}\left[\left(S^{(1)}(t) - S^{(2)}(t)\right)^{+}\right]. \tag{11}$$

Using $S^{(2)}$ as numeraire, we get that

$$V(0) = \mathbb{E}\left[S^{(2)}(t)\left(\frac{S^{(1)}(t)}{S^{(2)}(t)} - 1\right)\right] \tag{12}$$

$$= \mathbb{E}\left[S^{(2)}(t)\,(X(t) - 1)^{+}\right], \tag{13}$$

where $X(t) := S^{(1)}(t)/S^{(2)}(t)$.

Clearly, $\frac{S^{(2)}(t)}{S^{(2)}(0)}$ is a martingale, so we may define the Radon-Nikodym derivative

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_{\mathcal{F}_t} = \frac{S^{(2)}(t)}{S^{(2)}(0)} \tag{14}$$

from Girsanov's theorem we find that Q is an equivalent martingale measure, and that $W^{(2)}(t) = B^{(2)}(t) - \sigma_2 t$ is a Brownian motion under Q.

For $B^{(1)}$, we introduce a new P-Brownian motion $B^{(3)}$, which is independent of $B^{(1)}$ and define

$$B^{(1)}(t) = \rho B^{(2)}(t) + \sqrt{1-\rho^2}B^{(3)}(t). \tag{15}$$

From the measure change above we get that $W^1(t) = B^{(1)}(t) - \rho\sigma_2 t$ is a Q-Brownian motion.

Under Q, we find the dynamics of X is

$$X(t) = X(0)\exp\left(\sigma_1 W^{(1)}(t) - \sigma^2 W^{(2)}(t) - \frac{t}{2}\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2\rho\right)\right) \tag{16}$$

From (14) we find

$$V(0) = S^{(2)}(0)\,\mathbb{E}_Q\left[(X(t) - 1)^{+}\right]. \tag{17}$$

This combined with (16) shows that we have a call option on $X(t)$ with strike price 1 scaled by $S^{(2)}(0)$. Applying the Black-Scholes formula [14] then yields

$$V(0) = S^{(1)}(0)\Phi(d_1) - S^{(2)}(0)\Phi(d_2), \tag{18}$$

where

$$d_1 = \frac{\log\left(S^{(1)}(0)/S^{(2)}(0)\right) + \frac{1}{2}t\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma_2\rho\right)}{\sqrt{t\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma^2\rho\right)}}$$

$$d_2 = d_1 - \sqrt{t\left(\sigma_1^2 + \sigma_2^2 - \sigma_1\sigma^2\rho\right)},$$

which is what we wanted to prove. The quantities d_1, d_2 are found in the same way as they are found in the Black-Scholes formula. ■

5.2 Bachelier's model

In the Black-Scholes framework, the usual way of modelling the dynamics of an underlying asset is to model it as a geometric Brownian motion. The main reason for the choice of this model, is that it does not allow for negative prices, like an arithmetic Brownian motion may do. In this section, we will remove the positivity requirement on the spread, since the spread on two assets may be negative. We will therefore assume in this section that the dynamics of the spread on two assets $S_1(t)$ and $S_2(t)$, denoted $S(t)$ at time $t \geq 0$ is as follows:

$$\mathrm{d}S(t) = \mu S(t)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t).$$

However, the dynamics of the assets is modelled as geometric Brownian motions to avoid having a positive probability of negative prices. Hence,

$$\mathrm{d}S_i(t) = \mu S_i(t)\,\mathrm{d}t + \sigma_i S_i(t)\,\mathrm{d}W_i(t).$$

We then easily get from Itô's formula that

$$S_i(t) = S_i(0)\exp\left((\mu - \frac{1}{2}\sigma_i^2)t + \sigma_i W_i(t)\right).$$

We then have the following result which is **Proposition 2** in Carmona & Durrleman [17]. However, their proof is rather short, so we will here provide a complete proof.

Proposition 5.6. *If the value of the spread at maturity is assumed to have the Gaussian distribution, the price p of the call spread option with maturity T and strike K is given by*

$$p = \left(m(T) - Ke^{-rT}\right)\Phi\left(\frac{m(T) - Ke^{-rT}}{s(T)}\right) + s(T)\varphi\left(\frac{m(T) - Ke^{-rT}}{s(T)}\right), \tag{19}$$

where

$$m(T) = (x_2 - x_1)e^{(\mu - r)T}$$
$$s^2(T) = e^{2(\mu-r)T}\left(x_1^2\left(e^{\sigma_1^2 T} - 1\right) - 2x_1x_2\left(e^{\rho\sigma_1\sigma_2 T} - 1\right) + x_2^2\left(e^{\sigma_2^2 T} - 1\right)\right).$$

Proof: The idea of the proof is to approximate the distribution of $S(t)$ by the Gaussian distribution of the moments of $S_2(t) - S_1(t)$. This means that we should have

$$S(t) \sim N\left(\mathbb{E}[S_2(t) - S_1(t)], \mathrm{Var}[S_2(t) - S_1(t)]\right).$$

We can compute these quantities explicitly by moment matching.

$$\begin{aligned}
\mathbb{E}[S(t)] &= \mathbb{E}[S_2(t) - S_1(t)] \\
&= \sum_{i=1}^{2}(-1)^{i-1}\,\mathbb{E}\left[x_i \exp\left((\mu - \frac{1}{2}\sigma_i^2)t + \sigma_i W_i(t)\right)\right] \\
&= \sum_{i=1}^{2} x_i e^{(\mu - \frac{1}{2}\sigma_i^2)t} e^{\sigma_i\,\mathbb{E}[W_i(t)] + \frac{1}{2}\sigma_i^2 \mathrm{Var}[W_i(t)]} \\
&= (x_2 - x_1)e^{\mu t}
\end{aligned}$$

We can then multiply with the discount factor e^{-rt} and obtain $m(t)$.

In order to find $s^2(\cdot)$, we will first have to find the covariance of S_2 and S_1.

$$\begin{aligned}
\mathrm{Cov}(S_2(t), S_1(t)) &= \mathbb{E}[S_2(t)S_1(t)] - \mathbb{E}[S_2(t)]\,\mathbb{E}[S_1(t)] \\
&= \mathbb{E}\left[x_1x_2 \exp\left(\sum_{i=1}^{2}(\mu - \frac{1}{2}\sigma_i^2)t + \sigma_i W_i(t)\right)\right] - x_1x_2e^{2\mu t} \\
&= x_1x_2e^{2\mu t}\left(\mathbb{E}\left[e^{\sigma_2 W_2(t) - \frac{1}{2}\sigma_2^2 t + \sigma_1 W_1(t) - \frac{1}{2}\sigma_1^2 t}\right] - 1\right) \\
&= x_1x_2e^{2\mu t}\left(\exp\left(-\frac{1}{2}\sigma_2^2 t - \frac{1}{2}\sigma_1^2 t + \frac{1}{2}\sigma_2^2 t + \frac{1}{2}\sigma_1^2 t + \sigma_1\sigma_2 t\rho\right)\right) \\
&= x_1x_2e^{2\mu t}\left(e^{\sigma_1\sigma_2 t\rho} - 1\right).
\end{aligned}$$

Using this and the lemma above, we find that

$$\mathrm{Var}[S(t)] = \sum_{i=1}^{2} x_i^2 e^{2\mu t}\left(e^{\sigma_i^2 t} - 1\right) - 2x_1x_2e^{2\mu t}\left(e^{\sigma_1\sigma_2 t\rho} - 1\right),$$

and we may again multiply with the discount factor e^{-rt} to obtain the desired expression for $s^2(t)$.

We therefore find that in distribution we can write

$$e^{-rt}S(t) \overset{\mathrm{d}}{=} m(t) + s(t)Z,$$

where $Z \sim N(0,1)$. They payoff then becomes

$$\max\left(0, m(t) + s(t)Z - Ke^{-rt}\right),$$

which implies that

$$Z > \frac{Ke^{-rt} - m(t)}{s(t)} =: -d.$$

The price then becomes

$$\begin{aligned} p &= \int_{-d}^{\infty} \left(m(t) + s(t)x - Ke^{-rt}\right) \phi(x)\,\mathrm{d}x \\ &= \left(m(t) - Ke^{-rt}\right) \Phi(d) + \int_{-d}^{\infty} s(t)\phi(x)\,\mathrm{d}x \\ &= \left(m(t) - Ke^{-rt}\right) \Phi(d) + s(t)\phi(d), \end{aligned}$$

from which the result follows by substituting back for d. ■

Remark 5.7. When I first read the result above, it was a version from Rene Carmona's website [19], which I in error believed was the published version. In that version of the paper, the proof of the result above is wrong. Therefore, I corrected the proof and typed it here. Later, I found the printed article in SIAM Review, and found that the proof had been corrected. If the proof of the initial article had been correct, I would only have referenced the proof, but since my proof was already typed, I have decided to include the full proof including the intermediate steps that Carmona and Durrleman did not include in their paper.

5.3 Spread options on forward contracts under the framework of Heath, Jarrow and Morton

In the Heath, Jarrow & Morton framework we model directly on the spot. The arbitrage free forward dynamics is therefore

$$\frac{\mathrm{d}f(t,\tau)}{f(t,\tau)} = \sigma(t,\tau)\,\mathrm{d}B(t), \tag{20}$$

where $\sigma(\cdot,\tau) : [0,t] \to \mathbb{R}$ satisfies $\int_0^t |\sigma^2(s,\tau)|^2\,\mathrm{d}s < \infty$.

We can solve (20) by Itô's formula:

$$\begin{aligned}
\mathrm{d}\log f(t,\tau) &= f^{-1}(t,\tau)\,\mathrm{d}f(t,\tau) - \frac{1}{2} f^{-2}(t,\tau)(\,\mathrm{d}f(t,\tau))^2 \\
&= \sigma(t,\tau)\,\mathrm{d}B(t) - \frac{1}{2}\sigma^2(t,\tau)\,\mathrm{d}t.
\end{aligned}$$

We want to price an option, so from here on and out this subsection we have that $t \le T \le \tau$, where τ is the delivery time of the contract, T is the exercise time and t is the time the option is priced.

By continuing the calculation above, we find that

$$f(T,\tau) = f(t,\tau)\exp\left(\int_t^T \sigma(s,\tau)\,\mathrm{d}B(s) - \frac{1}{2}\int_t^T \sigma^2(s,\tau)\,\mathrm{d}s\right),$$

for $T \ge t$.

Theorem 5.8 ((†)). *The price $p(t)$ of a spread option at time t written on two forwards f_1 and f_2 is given by*

$$p(t) = e^{-r(T-t)}\left(f_1(t,\tau)\Phi(d_1) - f_2(t,\tau)\Phi(d_2)\right),$$

where

$$d_2 = \frac{\log(X(t)) - \frac{1}{2}\int_t^T \Sigma^2(s,\tau)\,\mathrm{d}s}{\sqrt{\int_t^T \Sigma^2(s,\tau)\,\mathrm{d}s}}$$

and

$$d_1 = d_2 + \sqrt{\int_t^T \Sigma^2(s,\tau)\,\mathrm{d}s},$$

and

$$\Sigma(t,\tau) = \left(\sigma_1^2(t,\tau) + \sigma_2^2(t,\tau) - 2\sigma_1(t,\tau)\sigma_2(t,\tau)\rho\right).$$

Proof:
We want to price a spread option on the spread of the two forward $f_1(t,T)$ and $f_2(t,T)$ with dynamics $\mathrm{d}f_i(t,T) = f_i(t,T)\sigma_i(t,T)\,\mathrm{d}B_i(t)$, so we get by Itô's formula that for $i = 1,2$

$$f_i(T,\tau) = f_i(t,\tau)\exp\left(\int_t^T \sigma_i(s,\tau)\,\mathrm{d}B_i(s) - \frac{1}{2}\int_t^T \sigma_i^2(s,\tau)\,\mathrm{d}s\right), \tag{21}$$

where $B_1(t)$ and $B_2(t)$ are two Brownian motions with (constant) correlation ρ.

We recall that the price of a call option written on a forward under the same circumstances with strike price K is given by the Black-76 formula [13], which states that

$$C(t, T, r, K) = e^{-r(T-t)} \left(f(0, T)\Phi(d_1) - K\Phi(d_2) \right), \tag{22}$$

where

$$d_1 = d_2 + \sqrt{\int_t^T \sigma^2(s, \tau)\, \mathrm{d}s}$$

and

$$d_2 = \frac{\log(f(t, \tau)/K) - \frac{1}{2}\int_t^T \sigma^2(s, \tau)\, \mathrm{d}s}{\sqrt{\int_t^T \sigma^2(s, \tau)\, \mathrm{d}s}}.$$

From equation (21) we see that the processes $T \mapsto f_i(T, \tau)$ are stochastic exponentials of the Doléans-Dade type, and are therefore martingales.[10]

From here on, we let $p(t)$ denote the price of a spread option written on two futures. It is defined as

$$p = \mathbb{E}\left[(f_1(T, \tau) - f_2(T, \tau))^+ \,\middle|\, \mathcal{F}_t \right],$$

where $(\cdot) = \max(0, \cdot)$ and $\mathcal{F}_t$ is a filtration. Further, we see that we may write

$$p = \mathbb{E}\left[f_2(T, \tau) \left(\frac{f_1(T, \tau)}{f_2(T, \tau)} - 1 \right)^+ \,\middle|\, \mathcal{F}_t \right].$$

Since $T \mapsto f_2(T, \tau)$ is a martingale, so is $\frac{f_2(T,\tau)}{f_2(t,\tau)}$, so we may define the following Radon-Nikodym derivative

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \frac{f_2(T, \tau)}{f_2(t, \tau)} = \mathcal{E}\left(\sigma(T, \tau) \cdot B_2(T) \right). \tag{23}$$

Then, by Girsanov's theorem

$$\mathrm{d}W_2(t) = \mathrm{d}B_2(t) - \sigma(t, \tau)\, \mathrm{d}t$$

is a Brownian motion under the measure Q. We will now define a new Brownian motion $B_3(t)$, independent of $B_2(t)$ such that

$$\mathrm{d}B_1(t) = \rho\, \mathrm{d}B_2(t) + \sqrt{1 - \rho^2}\, \mathrm{d}B_3(t).$$

[10] We point out that these stochastic exponentials usually are *local martingales*. However, since we assume that σ is deterministic (and therefore both adapted and independent of the Brownian motion), the aforementioned process is indeed a true martingale.

Further, under the measure change defined in (23), we find that under Q we get

$$\begin{aligned}
\mathrm{d}W_1(t) &= \rho\,\mathrm{d}W_2(t) + \sqrt{1-\rho^2}\,\mathrm{d}B_3(t) \\
&= \rho\,(\,\mathrm{d}B_2(t) - \sigma_2(t,\tau)\,\mathrm{d}t) + \sqrt{1-\rho^2}\,\mathrm{d}B_3(t) \\
&= \,\mathrm{d}B_1(t) - \rho\sigma_2(t,\tau).
\end{aligned}$$

To ease notation, we define

$$X(u) = \frac{f_1(u,\tau)}{f_2(u,\tau)}.$$

We can now find the dynamics of X under Q. Under P, we have

$$\begin{aligned}
X(T) &= \frac{f_1(t,\tau)}{f_2(t,\tau)} \exp\Bigg(\int_t^T \sigma_1(s,\tau)\,\mathrm{d}B_1(s) - \frac{1}{2}\int_t^T \sigma_1^2(s,\tau)\,\mathrm{d}s \\
&\qquad - \int_t^T \sigma_2(s,\tau)\,\mathrm{d}B_2(s) + \frac{1}{2}\int_t^T \sigma_2^2(s,\tau)\,\mathrm{d}s \Bigg) \\
&= X(t)\exp\Bigg(\int_t^T \sigma_1(s,\tau)\,\mathrm{d}B_1(s) - \frac{1}{2}\int_t^T \sigma_1^2(s,\tau)\,\mathrm{d}s \\
&\qquad - \int_t^T \sigma_2(s,\tau)\,\mathrm{d}B_2(s) + \frac{1}{2}\int_t^T \sigma_2^2(s,\tau)\,\mathrm{d}s \Bigg) \\
&= X(t)e^{Y(t,T)},
\end{aligned}$$

where

$$\begin{aligned}
Y(t,T) = \int_t^T \sigma_1(s,\tau)\,\mathrm{d}B_1(s) - \frac{1}{2}\int_t^T \sigma_1^2(s,\tau)\,\mathrm{d}s \\
- \int_t^T \sigma_2(s,\tau)\,\mathrm{d}B_2(s) + \frac{1}{2}\int_t^T \sigma_2^2(s,\tau)\,\mathrm{d}s.
\end{aligned}$$

Define $Y(t) = Y(0,t)$. From Itô's formula, we find that

$$\begin{aligned}
\mathrm{d}X(t) &= X(t)\,\mathrm{d}Y(t) + \frac{1}{2}X(t)\,(\,\mathrm{d}Y(t))^2 \\
&= X(t)\left(\sigma_1(t,\tau)\,\mathrm{d}B_1(t) - \frac{1}{2}\sigma_1^2(t,\tau)\,\mathrm{d}t - \sigma_2(t,\tau)\,\mathrm{d}B_2(t) + \frac{1}{2}\sigma_2^2(t,\tau)\,\mathrm{d}t\right) \\
&\quad + \frac{1}{2}X(t)\left(\sigma_1^2(t,\tau)\,\mathrm{d}t + \sigma_2^2(t,\tau)\,\mathrm{d}t - 2\rho\sigma_1(t,\tau)\sigma_2(t,\tau)\,\mathrm{d}t\right) \\
&= X(t)\left(\sigma_1(t,\tau)\,\mathrm{d}B_1(t) - \sigma_2(t,\tau)\,\mathrm{d}B_2(t) + \sigma_2^2(t,\tau)\,\mathrm{d}t - \rho\sigma_1(t,\tau)\sigma_2(t,\tau)\,\mathrm{d}t\right).
\end{aligned}$$

We can now find the dynamics under the measure Q.

$$\begin{aligned}
\mathrm{d}X(t) &= X(t)\left(\sigma_1(t,\tau)\left(\mathrm{d}W_1(t)+\rho\sigma_2(t,\tau)\,\mathrm{d}t\right)\right)\\
&\quad -\left(\sigma_1(t,\tau)\left(\mathrm{d}W_2(t)-\sigma_2(t,\tau)\,\mathrm{d}t\right)+\sigma_2^2(t,\tau)\,\mathrm{d}t-\rho\sigma_1(t,\tau)\sigma_2(t,\tau)\,\mathrm{d}t\right)\\
&= X(t)\left(\sigma_1(t,\tau)\,\mathrm{d}W_1(t)-\sigma_2(t,\tau)\,\mathrm{d}W_2(t)\right).
\end{aligned}$$

Hence

$$\frac{\mathrm{d}X(t)}{X(t)} = \sigma_1(t,\tau)\,\mathrm{d}W_1(t)-\sigma_2(t,\tau)\,\mathrm{d}W_2(t) \tag{24}$$

under the measure Q.

The bivariate SDE given in (24) is explicitly solvable using the two dimensional version of Itô's formula. We find that

$$\begin{aligned}
\mathrm{d}\log X(t) &= \frac{1}{X(t)}\,\mathrm{d}X(t)-\frac{1}{2}\frac{1}{X(t)^2}\left(\mathrm{d}X(t)\right)^2\\
&= \sigma_1(t,\tau)\,\mathrm{d}W_1(t)-\sigma_2(t,\tau)\,\mathrm{d}W_2(t)\\
&\qquad -\frac{1}{2}\underbrace{\left(\sigma_1^2(t,\tau)\,\mathrm{d}t+\sigma_2^2(t,\tau)\,\mathrm{d}t-2\sigma_1(t,\tau)\sigma_2(t,\tau)\rho\,\mathrm{d}t\right)}_{:=\Sigma(t,\tau)\,\mathrm{d}t}.
\end{aligned}$$

We can then integrate and exponentiate both sides, and find that

$$X(T) = X(t)\exp\left(\int_t^T \sigma_1(s,\tau)\,\mathrm{d}W_1(s)-\int_t^T \sigma_2(s,\tau)\,\mathrm{d}W_2(s)-\frac{1}{2}\int_t^T \Sigma(s,\tau)\,\mathrm{d}s\right).$$

Now, define a new Brownian motion under Q denoted $\tilde{W}(t)$. And we see that in distribution

$$X(T) = X(t)\exp\left(\int_t^T \Sigma(s,\tau)\,\mathrm{d}\tilde{W}(s)-\frac{1}{2}\int_t^T \Sigma^2(s,\tau)\,\mathrm{d}s\right), \tag{25}$$

which we recognize as a geometric Brownian motion of the same type as $f(T,\tau)$ defined in (20). And therefore we find, using the $\mathcal{F}_t$-measurability of $f(t,\tau)$ and the independent increment property of Brownian motion that

$$\begin{aligned}
p(t) &= e^{-r(T-t)}\,\mathbb{E}\left[f_2(T,\tau)\left(\frac{f_1(T,\tau)}{f_2(T,\tau)}-1\right)^+ \middle|\mathcal{F}_t\right]\\
&= e^{-r(T-t)}\,\mathbb{E}_Q\left[f_2(T,\tau)\frac{f_2(t,\tau)}{f_2(T,\tau)}\left(X(t)e^{\int_t^T \Sigma(s,\tau)\,\mathrm{d}\tilde{W}(s)-\frac{1}{2}\int_t^T \Sigma^2(s,\tau)\,\mathrm{d}s}-1\right)^+\right]\\
&= e^{-r(T-t)}f_2(t,\tau)\,\mathbb{E}_Q\left[(X(T)-1)^+\right],
\end{aligned}$$

which we recognize as a scaled version of the Black-76 formula with strike price $K = 1$. We therefore get

$$
\begin{aligned}
p(t) &\overset{B76}{=} e^{-r(T-t)} f_2(t,\tau) \left(\frac{f_1(t,\tau)}{f_2(t,\tau)} \Phi(d_1) - \Phi(d_2) \right) \\
&= e^{-r(T-t)} \left(f_1(t,\tau)\Phi(d_1) - f_2(t,\tau)\Phi(d_2) \right),
\end{aligned}
$$

which is what we wanted to show.
The quantities d_1 and d_2 are found the same way as in the Black-76 formula. ■

5.4 Pricing calendar spread options on forward contracts under the framework of Heath, Jarrow & Morton

In this section we will derive a result similar to the result in the previous section but for a calendar spread which is a spread on the same underlying asset with two different delivery times instead of a spread on two different underlyings.

Theorem 5.9 (†). *The price of a calendar spread option, $CS(t)$ written on a forward contract with fixed delivery times T and S is given by*

$$
CS = e^{-rt} \left(f(0,S)\Phi(d_1) - f(0,T)\Phi(d_2) \right), \tag{26}
$$

where d_1 and d_2 are given as

$$
\begin{aligned}
d_1 &= d_2 + \sqrt{\sum_{k=1}^{n} \int_0^t \left(\sigma_k(u,S) - \sigma_k(u,T) \right)^2 \, \mathrm{d}u} \\
d_2 &= \frac{\log \frac{f(0,S)}{f(0,T)} - \frac{1}{2} \sum_{k=1}^{n} \int_0^t \left(\sigma_k(u,S) - \sigma_k(u,T) \right)^2 \, \mathrm{d}u}{\sqrt{\sum_{k=1}^{n} \int_0^t \left(\sigma_k(u,S) - \sigma_k(u,T) \right)^2 \, \mathrm{d}u}}.
\end{aligned}
$$

It is assumed that $\int_0^t |\sigma_k^2(u,\cdot)| \, \mathrm{d}u < \infty$ for all $k = 1, 2, ..., n$.

Proof: We are in this case working with forward dynamics with an n-dimensional noise. We therefore have that

$$
\frac{\mathrm{d}f(t,T)}{f(t,T)} = \sum_{k=1}^{n} \sigma_k(t,T) \, \mathrm{d}B_k(t).
$$

Using the multidimensional Itô formula, see for example Chung [23] or Øksendal [62], we can state the explicit solution of the stochastic differential equation above

$$
f(t,T) = f(0,T) \exp \left(\sum_{k=1}^{n} \int_0^t \sigma_k(u,T) \, \mathrm{d}B_k(t) - \frac{1}{2} \sum_{k=1}^{n} \int_0^t \sigma_k^2(u,T) \, \mathrm{d}t \right),
$$

and we note that the result for $f(t, S)$ is similar.

Being Doléans Dade exponentials (where the integrands are deterministic), we see that both $t \mapsto f(t, T)$ and $t \mapsto f(t, S)$ are martingales. we find as in the previous subsection

$$\mathbb{E}\left[(f(t, S) - f(t, T))^+\right] = \mathbb{E}\left[f(t, T)\left(\frac{f(t, S)}{f(t, T)} - 1\right)^+\right].$$

We define

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \frac{f(t, T)}{f(0, T)} = \mathcal{E}\left(\sum_{k=1}^{n} \sigma_k(t, T) \cdot B_k(t)\right),$$

and by Girsanov's theorem

$$\mathrm{d}W_k(t) = \mathrm{d}B_k(t) - \sigma_k(t, T)$$

is a Brownian motion under the measure Q.
Under P we find that $\frac{f(t,S)}{f(t,T)}$ is given by

$$\begin{aligned}\frac{f(t, S)}{f(t, T)} = \frac{f(0, S)}{f(0, T)} \exp\Bigg(&\sum_{k=1}^{n} \int_0^t (\sigma(u, S) - \sigma(u, T))\, \mathrm{d}B(u) \\ &- \frac{1}{2}\sum_{k=1}^{n}\left(\int_0^t \sigma^2(u, S) - \sigma^2(u, T)\right)\, \mathrm{d}u\Bigg),\end{aligned}$$

We define $X(t) := \frac{f(t,S)}{f(t,T)}$, and by Itô's formula, we find that the dynamics of $X(t)$ are

$$\begin{aligned}\mathrm{d}X(t) &= X(t)\left(\sum_{k=1}^{n}(\sigma_k(t, S) - \sigma_k(t, T))\, \mathrm{d}B_k(t) - \frac{1}{2}\sum_{k=1}^{n}(\sigma_k^2(t, S) - \sigma_k^2(t, T))\, \mathrm{d}t\right) \\ &\quad + \frac{1}{2}\left(\sum_{k=1}^{n}(\sigma_k^2(t, S) + \sigma_k^2(t, T) - 2\sigma_k(t, S)\sigma_k(t, T))\, \mathrm{d}t\right) \\ &= X(t)\left(\sum_{k=1}^{n}(\sigma_k(t, S) - \sigma_k(t, T))\, \mathrm{d}B_k(t) + \sigma_k^2(t, T)\, \mathrm{d}t - \sigma_k(t, S)\sigma_k(t, T)\, \mathrm{d}t\right),\end{aligned}$$

which under the measure Q becomes

$$\begin{aligned}
\mathrm{d}X(t) &= X(t)\sum_{k=1}^{n}\Big((\sigma_k(t,S)-\sigma_k(t,T))(\,\mathrm{d}W_k(t)+\sigma_k(t,T)\,\mathrm{d}t) \\
&\qquad + \sigma_k^2(t,T)\,\mathrm{d}t - \sigma_k(t,S)\sigma_k(t,T)\,\mathrm{d}t\Big) \\
&= X(t)\sum_{k=1}^{n}((\sigma_k(t,S)-\sigma_k(t,T))\,\mathrm{d}W_k(t))\,,
\end{aligned}$$

which we recognize as a geometric Brownian motion, which by Itô's formula has solution

$$\begin{aligned}
X(t) = X(0)\exp\Bigg(&\sum_{k=1}^{n}\int_0^t(\sigma_k(u,S)-\sigma_k(u,T))\,\mathrm{d}W_k(u) \\
&-\frac{1}{2}\sum_{k=1}^{n}\int_0^t(\sigma_k(u,S)-\sigma_k(u,T))^2\,\mathrm{d}u\Bigg).
\end{aligned}$$

Hence the price of the calendar spread is given as

$$\begin{aligned}
CS &= e^{-rt}\,\mathbb{E}\left[f(t,T)(X(t)-1)^+\right] \\
&= e^{-rt}f(0,T)\left(\frac{f(0,S)}{f(0,T)}\Phi(d_1)-\Phi(d_2)\right)
\end{aligned}$$

and finally

$$CS = e^{-rt}\left(f(0,S)\Phi(d_1)-f(0,T)\Phi(d_2)\right), \tag{27}$$

which is what we wanted to show. For the last part we used the same results as in the earlier proofs.

The quantities d_1 and d_2 are found as in the Black-76 formula, see for example Black [13] or Benth, Benth & Koekebakker [10]. ■

This theorem may also be proved using the approximation result presented in section 5.6.

5.5 Spread options when the underlying asset follows an Arithmetic Brownian motion

We will in this section derive similar results when the dynamics of the underlying follows an arithmetic Brownian motion. That is,

$$\mathrm{d}X(t) = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}B(t).$$

Due to the model's simple nature, we will in this section consider the case where both the mean μ and volatility σ are time dependent. We will prove pricing results for both spread options of the Margrabe type and calendar spread options.

Theorem 5.10 (†). *The price of a Margrabe style spread option where the underlying assets follow an arithmetic Brownian motion, is given as*

$$p = e^{-rt}\left(m(t)\Phi(d) + \sqrt{\tilde{\sigma}}(t)\phi(d)\right),$$

where

$$d = \frac{\int_0^t \mu_1(u) - \mu_2(u)\,\mathrm{d}u}{\sqrt{\tilde{\sigma}(u)\,\mathrm{d}u}},$$

and $m(t) = \int_0^t \mu_1(s) - \mu_2(s)\,\mathrm{d}s$, $\tilde{\sigma}(t) = \sigma_1^2(t) + \sigma_2^2(t) - 2\sigma_1(t)\sigma_2(t)\rho$, $\phi(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}}$ *and* Φ *denotes the standard normal cumulative distribution function.*

Proof: The dynamics for the two assets is given as

$$\mathrm{d}S_i(t) = \mu_i(t)\,\mathrm{d}t + \sigma_i(t)\,\mathrm{d}B_i(t), \tag{28}$$

Where $B_1(t)$ and $B_2(t)$ are two Brownian motions with correlation ρ.
The SDE given in (28) has the obvious solution

$$S_i(t) = \int_0^t \mu_i(u)\,\mathrm{d}u + \int_0^t \sigma_i(u)\,\mathrm{d}B_i(u).$$

We can then as above compute the price

$$\begin{aligned} p =& \ \mathbb{E}\left[(S_1(t) - S_2(t))^+\right] \\ =& \ \mathbb{E}\left[\left(\int_0^t \mu_1(u) - \mu_2(u)\,\mathrm{d}u + \int_0^t \sigma_1(u)\,\mathrm{d}B_1(u) - \int_0^t \sigma_2(u)\,\mathrm{d}B_2(u)\right)^+\right] \\ =& \ \mathbb{E}\left[\left(m(t) + \int_0^t \sigma_1(u)\,\mathrm{d}B_1(u) - \int_0^t \sigma_2(u)\,\mathrm{d}B_2(u)\right)^+\right]. \end{aligned}$$

In distribution we have that

$$\begin{aligned} & m(t) + \int_0^t \sigma_1(u)\,\mathrm{d}B_1(u) - \int_0^t \sigma_2(u)\,\mathrm{d}B_2(u) \\ & \stackrel{\text{dist}}{=} m(t) + \left(\int_0^t \sigma_1^2(u) + \sigma_2^2(u) - 2\sigma_1(u)\sigma_2(u)\rho\,\mathrm{d}u\right)^{1/2} Y \\ & = m(t) + \sqrt{\int_0^t \tilde{\sigma}(u)\,\mathrm{d}uY}, \end{aligned}$$

where $Y \sim N(0,1)$.
Hence, if we require that

$$Y > -\frac{m(t)}{\sqrt{\int_0^t \tilde{\sigma}(u)\,\mathrm{d}u}} =: -d,$$

we then get that

$$\begin{aligned} p &= \frac{1}{\sqrt{2\pi}} \int_{-d}^{\infty} \left(m(t) + \sqrt{\tilde{\sigma}(t)} y \right) e^{-y^2/2}\,\mathrm{d}y \\ &= m(t)\Phi(d) + \sqrt{\tilde{\sigma}}(t)\phi(d). \end{aligned}$$

Applying the discount factor yields the desired result.

■

We can now turn our attention to a calendar spread option where the underlying forward contract follows an arithmetic model.

Theorem 5.11 (†). *The price of a calendar spread option when the underlying has delivery times S and T and follows an arithmetic Brownian motion is given as*

$$p = e^{-rt} \left(\int_0^t m(u,S,T)\,\mathrm{d}u \Phi(d) + \tilde{\sigma}(t)\phi(d) \right),$$

where

$$d = \frac{\int_0^t m(u,S,T)\,\mathrm{d}u}{\tilde{\sigma}(t,S,T)}.$$

In d, $m(t,T,S) = \mu(t,S) - \mu(t,T)$ and $\tilde{\sigma}(t,S,T) = \sqrt{\int_0^t (\sigma(u,S) - \sigma(u,T))^2\,\mathrm{d}u}$.

Proof: The SDE governing the forward contract is now given as

$$\mathrm{d}f(t,T) = \mu(t,T)\,\mathrm{d}t + \sigma(t,T)\,\mathrm{d}B(t).$$

Hence

$$\begin{aligned} p &= e^{-rt}\,\mathbb{E}\left[(f(t,S) - f(t,T))^+ \right] \\ &= e^{-rt}\,\mathbb{E}\left[\left(\int_0^t \mu(u,S) - \mu(u,T)\,\mathrm{d}u + \int_0^t \sigma(u,S) - \sigma(u,T)\,\mathrm{d}B(u) \right)^+ \right] \\ &=: e^{-rt}\,\mathbb{E}\left[\left(\int_0^t m(u,T,S)\,\mathrm{d}u + \int_0^t \Sigma(u,T,S)\,\mathrm{d}B(u) \right)^+ \right]. \end{aligned}$$

In distribution we have that

$$\int_0^t m(u,S,T)\mathrm{d}u+\int_0^t \Sigma(u,S,T)\mathrm{d}B(u) = \int_0^t m(u,S,T)\mathrm{d}u+\sqrt{\int_0^t \Sigma^2(u,S,T)\,\mathrm{d}u}Y,$$

where $Y \sim N(0,1)$. Hence if we define $\sqrt{\int_0^t \Sigma^2(u,S,T)\,\mathrm{d}u} = \tilde{\sigma}(t,S,T)$ and require that $Y > -\frac{\int_0^t m(u,S,T)}{\tilde{\sigma}(t,S,T)} =: -d$, then

$$p = e^{-rt}\left(\int_0^t m(u)\,\mathrm{d}u\Phi(d) + \tilde{\sigma}(t)\phi(d)\right),$$

which is what we wanted to prove.

■

5.6 An approximation for options on the spread of geometric Brownian motions

In this section we assume that the risk neutral price dynamics of the assets are given by the following stochastic differential equations

$$\begin{aligned}\mathrm{d}S_1(t) &= S_1(t)[(r-q_1)\,\mathrm{d}t + \sigma_1\,\mathrm{d}W_1(t)]\\ \mathrm{d}S_2(t) &= S_2(t)[(r-q_2)\,\mathrm{d}t + \sigma_2\,\mathrm{d}W_2(t)],\end{aligned}$$

where q_1 and q_2 are the instantaneous dividend yields, the volatilities $\sigma_1, \sigma_2 \in \mathbb{R}^+$ and W_1 and W_2 are Brownian motions with correlation ρ.

By appealing to earlier arguments (or to Carmona & Durrleman [17]), we see that a call option on the spread of S_1 and S_2 with strike price K can be written as

$$p = e^{-rT}\mathbb{E}\left[\max\left(0, x_2 e^{(r-q_2-\sigma_2^2/2)T+\sigma_2 W_2(T)} - x_1 e^{(r-q_1-\sigma_1^2/2)T+\sigma_1 W_1(t)} - K\right)\right]. \tag{29}$$

Carmona & Durrleman introduce the following notation for the expectation that need to be calculated in order to price the option:

$$\Pi = \Pi(\alpha,\beta,\gamma,\delta,\kappa,\rho) = \mathbb{E}\left[\left(\alpha e^{\beta X_1-\beta^2/2} - \gamma e^{\delta X_2-\delta^2/2} - \kappa\right)^+\right], \tag{30}$$

where $\alpha, \beta, \gamma, \delta$ and κ are real valued constants, X_1 and X_2 are jointly standard normal random variables with correlation ρ. We see for example that (29) and

(30) coincide in distribution whenever $\alpha = x_2 e^{-q_2 T}$, $\beta = \sigma_2\sqrt{T}$, $\gamma = x_1 e^{-q_1 T}$, $\delta = \sigma_1\sqrt{T}$, $\kappa = K$ and multiply with the discount factor e^{-rT}.
We can now state the main result from Carmona & Durrleman [17]. The following theorem is a mashup of **Proposition 6.1** and **Proposition 6.2** in Carmona & Durrleman [17].

Theorem 5.12. *Define*

$$\begin{aligned} d^* = {} & \frac{1}{\sigma\cos(\theta^* - \psi)\sqrt{T}} \log\left(\frac{x_2 e^{-q_2 T}\sigma_2 \sin(\theta^* + \phi)}{x_1 e^{-q_1 T}\sigma_1 \sin\theta^*}\right) \\ & - \frac{1}{2}(\sigma_2\cos(\theta^* + \phi) + \sigma_1\cos\theta^*)\sqrt{T}, \end{aligned}$$

*where the angles ϕ and ψ are chosen such that $\cos\phi = \rho$ and $\cos\psi = \frac{\sigma_1 - \rho\sigma_2}{\sigma}$ and θ^* is the solution to equation **(37)** in Carmona & Durrleman [17]. Then*

$$\begin{aligned} \hat{p} = {} & x_2 e^{-q_2} T \Phi\left(d^* + \sigma_2\cos(\theta^* + \phi)\sqrt{T}\right) \\ & - x_1 e^{-q_1 T}\Phi\left(d^* + \sigma_1 \sin\theta^*\sqrt{T}\right) - K e^{-rT}\Phi\left(d^*\right). \end{aligned} \tag{31}$$

Moreover, the approximation $\hat{p}$ is equal to the true price p when $K = 0$, or $x_1 = 0$, or $x_2 = 0$, or $\rho = -1$. In particular, we retrieve Margrabe's formula whenever $K = 0$ and the classical Black-Scholes formula whenever $x_1 = 0$ or $x_2 = 0$.

Proof: See Carmona & Durrleman [17] and the references therein. ■

Carmona & Durrleman states in the same paper a way to find the price of a calendar spread option, but without proof. We will therefore prove it here.

Corollary 5.13. *Let $f(t,T)$ and $f(t,S)$ denote a forward contract with delivery times T and S. The coefficients from* (30) *are then given by*

$$\begin{aligned} \alpha &= f(t,T) \\ \beta &= \sqrt{\int_0^t \sigma^2(u,T)\,\mathrm{d}u} \\ \gamma &= f(t,S) \\ \delta &= \sqrt{\int_0^t \sigma^2(u,S)\,\mathrm{d}u} \\ \rho &= \frac{1}{\beta\delta}\int_0^t \sigma(u,T)\sigma(u,S)\,\mathrm{d}u. \end{aligned}$$

The price is then given by

$$CS = e^{-rt}\left(f(0,S)\Phi(d_1) - f(0,T)\Phi(d_2)\right). \tag{32}$$

Proof: The formulas for α, β, γ and δ follows from the representation of (30). The correlation coefficient ρ is computed in the following manner:

We know that $xy = \frac{1}{4}\left((x+y)^2 - (x-y)^2\right)$. So we get that

$$\begin{aligned}\int_0^t \sigma(u,S)\,\mathrm{d}W(u) \int_0^t \sigma(u,T)\,\mathrm{d}W(u) = \frac{1}{4}\Bigg(&\left(\int_0^t (\sigma(u,S)+\sigma(u,T))\,\mathrm{d}W(u)\right)^2 \\ &- \left(\int_0^t (\sigma(u,S)-\sigma(u,T))\,\mathrm{d}W(u)\right)^2\Bigg).\end{aligned}$$

By applying the Itô isometry, we can calculate the covariance

$$\begin{aligned}&\mathrm{Cov}\left(\int_0^t \sigma(u,S)\,\mathrm{d}W(u), \int_0^t \sigma(u,T)\,\mathrm{d}W(u)\right) \\ &= \mathbb{E}\left[\int_0^t \sigma(u,S)\,\mathrm{d}W(u) \int_0^t \sigma(u,T)\,\mathrm{d}W(u)\right] \\ &= \mathbb{E}\left[\frac{1}{4}\left(\left(\int_0^t (\sigma(u,S)+\sigma(u,T))\,\mathrm{d}W(u)\right)^2 - \left(\int_0^t (\sigma(u,S)-\sigma(u,T))\,\mathrm{d}W(u)\right)^2\right)\right] \\ &= \frac{1}{4}\Bigg(\int_0^t \sigma^2(u,S) + \sigma^2(u,T) + 2\sigma(u,S)\sigma(u,T)\,\mathrm{d}u \\ &\qquad - \int_0^t \sigma^2(u,S) + \sigma^2(u,T) - 2\sigma(u,S)\sigma(u,T)\,\mathrm{d}u\Bigg) \\ &= \int_0^t \sigma(u,T)\sigma(u,S)\,\mathrm{d}u,\end{aligned}$$

from which the correlation coefficient ρ follows. ■

Remark 5.14. Using this result we can find the price of the spread option described in Theorem 5.9. The price of the option follows from the fact that the strike price K is zero, hence by the above theorem above equation (30) reduces to Margrabe's formula, and we refer to Theorem 5.9 for that calculation.

Remark 5.15. In addition to the approximation given by Carmona & Durrleman and Bachelier's model, there is at least one more approximation result for spread options worth mentioning, and that is Kirk's formula proposed by Kirk in [48].

However, it is shown in Carmona & Durrleman [17] that their method performs just as well as Kirk's approximation, and in some cases better. Moreover, computing the greeks is easier as well, leading to more sensible hedging portfolios, in contrast to Kirk's approximation where one ends up with two delta hedges. We refer to Carmona & Durrleman [17] for further discussion on this topic.

6 Ornstein-Uhlenbeck process

We will now discuss an example where we do not have the stochastic process $X(t)$ in all terms. This example is often called the Ornstein-Uhlenbeck process or the Vasicek model. The Ornstein-Uhlenbeck process is a particulary interesting case for finance, since it is used as a basic tool for modelling interest rates and asset prices, see e.g Benth [9, p.43]. It is also, due to its natural way of reversion to the mean, used to model the spot price in electricity and related markets. Moreover, this equation also approximately describes a one dimensional Brownian motion of a free particle in fluid.

When modelling spot prices for electricity markets, one must allow for jumps. This condition is easily enforced in the class of Ornstein-Uhlenbeck processes and is done by allowing the stochastic driver to allow for jumps, usually letting the stochastic driver be a compound Poisson process. However, the derivation of the solution of an Ornstein-Uhlenbeck process are the same as in the Gaussian case, so in this chapter, we confine ourselves to only work within the framework of Brownian motion-driven Ornstein-Uhlenbeck processes.

6.1 The finite dimensional case

The Ornstein-Uhlenbeck stochastic differential equation is defined on $\mathbb{R}$ to be

$$\mathrm{d}X_t = -\alpha X(t)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t), \tag{33}$$

where $\alpha, \sigma \in \mathbb{R}_+$ are given constants, $X(0) \in \mathbb{R}$. We find that the integrating factor is $e^{\alpha t}$, and we define $g(t,x) = xe^{\alpha t}$. The following result is well known, and is included for completion.

Proposition 6.1. *The solution of* (33) *is*

$$X(t) = X(0)e^{-\alpha t} + \int_0^t \sigma e^{-\alpha(t-s)}\,\mathrm{d}B(s). \tag{34}$$

Proof: By Itô's formula we get

$$\mathrm{d}g(t, X(t)) = \mathrm{d}xe^{\alpha t} = \alpha X(t)e^{\alpha t}\,\mathrm{d}t + e^{\alpha t}\,\mathrm{d}X(t).$$

Consequently,

$$e^{\alpha}\,\mathrm{d}X(t) = \mathrm{d}X(t)e^{\alpha t} - \alpha X(t)e^{\alpha t}. \tag{35}$$

We then multiply (33) with the integrating factor, we have

$$e^{\alpha t}\,\mathrm{d}X(t) = -\alpha X(t)e^{\alpha t}\,\mathrm{d}t + \sigma e^{\alpha t}\,\mathrm{d}B(t). \tag{36}$$

Comparing (35) and (36) yields

$$\begin{aligned} e^{\alpha t}\,\mathrm{d}X(t) &= -\alpha X(t)e^{\alpha t}\,\mathrm{d}t + \sigma e^{\alpha t}\,\mathrm{d}B(t) \\ &= \mathrm{d}e^{\alpha t}X(t) - \alpha X(t)e^{\alpha t}. \end{aligned}$$

This implies that

$$e^{\alpha t}\,\mathrm{d}X(t) = \sigma e^{\alpha t}\,\mathrm{d}B(t) = \mathrm{d}e^{\alpha t}X(t).$$

Integrating from 0 to t, we obtain

$$e^{\alpha t}X(t) = X(0) + \sigma\int_0^t e^{\alpha s}\,\mathrm{d}B(s)$$

Rearranging terms shows us that

$$X(t) = X(0)e^{-\alpha t} + \sigma e^{-\alpha t}\int_0^t e^{\alpha s}\,\mathrm{d}B(s),$$

which is what we wanted to show. ■

From this we can easily derive its mean and variance. The mean is clearly given as $\mathbb{E}X(t) = X(0)e^{-\alpha t}$, since the stochastic integral has expectation zero. The variance is given by the Itô isometry and we find that it is

$$\begin{aligned} \mathrm{Var}[X(t)] &= \int_0^t \sigma^2 e^{-2\alpha(t-s)}\,\mathrm{d}s \\ &= \sigma^2 e^{-2\alpha t}\frac{e^{2\alpha t}-1}{2\alpha} = \frac{\sigma^2}{2\alpha}(1-e^{-2\alpha t}) \end{aligned}$$

. We therefore see that for the Ornstein-Uhlenbeck process there exists a limiting distribution. We therefore find

Corollary 6.2. *As $t \to \infty$, the Ornstein-Uhlenbeck process is normally distributed with mean* 0 *and variance $\sigma^2/2\alpha$.*

Remark 6.3. By slightly modifying **Lemma 10.1** in Benth, Benth & Koekebakker [10], we find that if $X(t)$ is an $\mathbb{R}^n$-valued process, its solution is given as

$$\mathbf{X}(t) = \mathbf{X}(0)e^{tA} + \int_0^t e^{A(t-u)}\mathbf{e}_p\sigma\,\mathrm{d}B(u),$$

where A is an $n\times n$ matrix whose eigenvalues have negative real part (See Benth, Benth & Koekebakker chapter 10 for details). We know from **Lemma 2.10** in Meiss [56] that if A has eigenvalues with negative real part, then $e^{tA} \to 0$ as $t \to \infty$. We therefore see that the same mean reverting properties hold as in the 1-dimensional case.

6.2 The infinite dimensional case

Like in the finite dimensional case, we define the stochastic differential equation bearing the name of Ornstein-Uhlenbeck:

$$\mathrm{d}X(t) = (AX(t) + f(t))\ \mathrm{d}t + B\ \mathrm{d}W(t),\ X(0) = x, \tag{37}$$

where $A : D(A) \subset H \to H$, $B : U \to H$, $f \in H$ and W is an H-valued Wiener process and $D(A)$ denotes the domain of A. U and H are separable Hilbert spaces.

According to chapter 5 in Da Prato & Zabczyk, we define the following three hypotheses

(H1) A generates a C_0 semigroup $S(\cdot)$ in H, and $B \in L(U, H)$

(H2) f is predictable with Bochner integrable trajectories on an arbitrary finite interval $[0, T]$, and x is $\mathcal{F}_0$-measurable

(H3) $\int_0^T \|S(r)B\|_{L_2^0}^2\,\mathrm{d}r = \int_0^T \mathrm{Tr}(S(r)BQB^*S^*(r))\,\mathrm{d}r < \infty$.

Define the *stochastic convolution* as

$$W_A(t) = \int_0^t S(t-s)B\ \mathrm{d}W(s). \tag{38}$$

We refer to **Theorem 5.10** in Da Prato & Zabczyk for other representations of (38).

It is then proven in **Theorem 5.4** and **Theorem 5.11** in Da Prato & Zabczyk [65] that if the three aforementioned hyotheses (H1), (H2) and (H3) hold, then there exists a unique weak solution $X(t)$ to (37). If there exists $0 < \alpha < \frac{1}{2}$ such that

$$\int_0^T t^{-2\alpha}\|S(r)B\|_{L_2^0}^2\,\mathrm{d}r < \infty,$$

then $X(t)$ admits a continuous version. $X(t)$ is given as

$$X(t) = S(t)x + \int_0^t S(t-s)f(s)\,\mathrm{d}s + \int_0^t S(t-s)B\ \mathrm{d}W(s). \tag{39}$$

Remark 6.4. The solution $X(t)$ given in (39), is completely analogous to the solutions we may find in the finite dimensional case. If one for example wants to model continuous time autoregressive processes, then the stochastic differential equation takes on the same form as in (37), but where A is an $n \times n$ matrix whose eigenvalues have negative real part[11], given by **(10.12)** in Benth, Benth & Koekebakker [10], $B = \mathbf{e}_n\sigma(t)$ and $X(t)$ is a process in $\mathbb{R}^n$. We therefore have

$$\mathrm{d}X(t) = (A\mathbf{X}(t) + f(t))\ \mathrm{d}t + \mathbf{e}_n\sigma(t)\,\mathrm{d}W(t),\ X(0) = x.$$

Using variation of constants and Itô's formula, we find the solution to (6.4)

$$X(t) = e^{tA}x + \int_0^t e^{(t-s)A} f(s)\,\mathrm{d}s + \int_0^t e^{(t-s)A}\mathbf{e}_n\sigma(s)\,\mathrm{d}W(s).$$

We therefore see that the semigroup $S(t)$ in this case is of the form

$$S(t) = e^{tA} := \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}.$$

If $H = \mathbb{R}$, then $A = -\alpha$, where $\alpha \in \mathbb{R}_+ - \{0\}$, hence $S(t) = e^{-\alpha t}$. We refer to the previous subsection or any book on mathematical finance for a discussion about the distributional properties of Ornstein-Uhlenbeck processes in finite dimensional spaces.

In the infinite dimensional case, we see that

$$\mathbb{E}[X(t)] = S(t)x + \int_0^t S(t-s)f(s)\,\mathrm{d}s$$

since $\mathbb{E}[W(t)] = 0$, and that

$$\mathrm{Var}[X(t)] = \int_0^t (S(t-s)B)^2\,\mathrm{d}s,$$

by the Itô isometry.

It is possible to extend the Ornstein-Uhlenbeck processes to be driven by general Lévy processes. Using γ as characteristic exponent, one can show that

$$\mathbb{E}\left[e^{i\langle u,Y(t)\rangle}\right] = \exp\left(\gamma(S^*(t)(u)) + \int_0^t \gamma(S^*(r)(u))\,\mathrm{d}r\right),$$

see for example Applebaum [1] and the references therein.

[11] The fact that the eigenvalues have negative real part ensures the existence of a limiting distribution, since $e^{tA} \to 0$ as $t \to \infty$ if and only if the eigenvalues have negative real part. See for example **Lemma 2.10** in Meiss [56].

7 Stochastic modelling of electricity markets

In this section we will give some results on the stochastic modelling of electricity prices.

We basically have two models for the spot price: arithmetic models and geometric models. If we let $S(t)$ denote the spot price at time t, then a arithmetic model is defined as:

$$S(t) = \Lambda(t) + \sum_{i=1}^{m} X_i(t) + \sum_{j=1}^{n} Y_j(t).$$

Similarly, a geometric spot price model is defined as

$$S(t) = \exp\left(\Lambda(t) + \sum_{i=1}^{m} X_i(t) + \sum_{j=1}^{n} Y_i(t)\right).$$

In both cases, Λ is some seasonality function, which we in subsequent chapters will assume is constantly zero. The functions $X(t)$ and $Y(t)$ are Ornstein-Uhlenbeck processes of the form

$$\mathrm{d}X(t) = -\alpha X(t)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t)$$

and

$$\mathrm{d}Y(t) = -\beta Y(t)\,\mathrm{d}t + \eta\,\mathrm{d}I(t),$$

where $B(t)$ is a Brownian motion and $I(t)$ is a compound Poisson process. By following the previous chapter, it can be shows that the solutions are given as

$$X(t) = e^{-\alpha t}X(0) + \int_0^t \sigma e^{-\alpha(t-s)}\,\mathrm{d}B(s)$$

and

$$Y(t) = e^{-\beta t}Y(0) + \int_0^t \eta e^{-\beta(t-s)}\,\mathrm{d}I(s).$$

We will now provide an example of a geometric model, where $\Lambda(t) = 0$ and $Y(t) = 0$ for all t. This model is known as a simple case of the *Schwartz model*, and was introduced by Schwartz [72] in 1997. This model is, according to Benth, Benth & Koekebakker [10] the classical model for commodity markets, as it is an extension of the geometric Brownian motion allowing for mean reversion.

Example 7.1 (Pricing a forward contract using the Schwartz model)**.** Let S be a commodity modelled by a simple Schwartz model. That is

$$S(t) = S(0)e^{X(t)},$$

where
$$\mathrm{d}X(t) = -\alpha X(t)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t).$$
As will be shown in the coming subsection, the price at time t of a forward contract, denoted $f(t, T)$, is given as
$$f(t, T) = \mathbb{E}\left[S(T)|\mathcal{F}_t\right].$$
We want to compute $f(t, T)$ and find a stochastic differential equation for f.

Using that we have found the solution of $X(\cdot)$, we find
$$\begin{aligned}
\mathbb{E}[S(T)|\mathcal{F}_t] &= S(0)\exp\left(e^{-\alpha(T-t)}X(t)\right)\,\mathbb{E}\left[\exp\left(\int_t^T \sigma e^{-\alpha(T-s)}\,\mathrm{d}B(s)\right)|\mathcal{F}_t\right] \\
&= S(0)\exp\left(e^{-\alpha(T-t)}X(t) + \frac{1}{2}\int_t^T \sigma^2 e^{-2\alpha(T-s)}\,\mathrm{d}s\right) \\
&= S(0)\exp\left(e^{-\alpha(T-t)}X(t) + \frac{\sigma^2}{4\alpha}\left(1 - e^{-2\alpha(T-t)}\right)\right).
\end{aligned}$$
A simple application of Itô's formula reveals that
$$\mathrm{d}f(t, T) = f(t, T)\sigma e^{-\alpha(T-t)}\,\mathrm{d}B(t),$$
which shows that $f(t, T)$ is a driftless geometric Brownian motion. Hence we may rewrite the expression as
$$f(t, T) = f(0, T)\exp\left(\int_0^t \sigma e^{-\alpha(T-t)}\,\mathrm{d}B(s) - \frac{1}{2}\int_0^t \sigma^2 e^{-2\alpha(T-s)}\,\mathrm{d}s\right).$$
□

7.1 Forward pricing

Assume we want to buy a contract delivering some commodity at a fixed time, we then need to ask ourselves, what is the fair price of such a contract? Imagine that we sell the commodity on the spot and buy the contract. We therefore get that the payoff is $S(\tau) - f(t, \tau)$. By a standard no-arbitrage argument we get that the value of the contract is
$$C(t, \tau) = \mathbb{E}_Q\left[e^{-r(\tau-t)}\left(S(\tau) - f(t, \tau)\right)|\mathcal{F}_t\right],$$
for some pricing probability $Q \sim P$.

However, it does not cost anything to enter such a contract, and since $f(t,\tau)$ is $\mathcal{F}_t$-measurable, we get that

$$f(t,\tau) = \mathbb{E}_Q\left[S(\tau)|\mathcal{F}_t\right]. \tag{40}$$

The above argument holds for all fixed delivery contracts. However, some commodities, like power, is delivered over a period $[\tau_1, \tau_2]$. The pricing of such contracts is a little more involved. We know that the payoff from such a position will is

$$\int_{\tau_1}^{\tau_2} e^{-r(u-t)}\left(S(u) - F(t,\tau_1,\tau_2)\right)\,\mathrm{d}u.$$

The entrance cost is still zero, so we get by a similar argument that

$$F(t,\tau_1,\tau_2)\int_{\tau_1}^{\tau_2} e^{-ru}\,\mathrm{d}u = \mathbb{E}_Q\left[\int_{\tau_1}^{\tau_2} e^{-ru}S(u)\,\mathrm{d}u|\mathcal{F}_t\right].$$

We therefore get that the price is

$$F\left(t,\tau_1,\tau_2\right) = \int_{\tau_1}^{\tau_2}\frac{e^{-ru}}{\int_{\tau_1}^{\tau_2} e^{-rv}\,\mathrm{d}v}f(t,u)\,\mathrm{d}u.$$

We know that the exponential function if defined as

$$e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!},$$

which shows that

$$e^{-ru} = 1 - ru + \frac{(ru)^2}{2} - \frac{(ru)^3}{3!} + ...$$

And therefore, whenever r and u are adequately small, we get that $e^{-ru} \approx 1$. Then we find that

$$F\left(t,\tau_1,\tau_2\right) = \frac{1}{\tau_2-\tau_1}\int_{\tau_1}^{\tau_2} f(t,u)\,\mathrm{d}u.$$

When we are close to delivery, we can show that the forward price $f(t,T)$ for a fixed delivery contract tends towards the spot price. This is shown easily by $f(T,T) = \mathbb{E}[S(T)|\mathcal{F}_T] = S(T)$. However, we see that

$$\lim_{t\to\tau_1} F\left(t,\tau_1,\tau_2\right) \neq S(\tau_1) \tag{41}$$

with probability 1. But, if we shrink the delivery period $[\tau_1, \tau_2]$, meaning that $\tau_2 \to \tau_1$, it is shown in **Proposition 4.3** in Benth, Benth & Koekebakker [10] that the price of the forward contract tends to the forward price with delivery at τ_1. Meaning that

$$\lim_{\tau_2\to\tau_1} F(t,\tau_1,\tau_2) = f(t,\tau_1).$$

(41) follows from **Proposition 4.2** in Benth, Benth & Koekebakker.

7.2 The Musiela parametrization

In this section we will introduce a very convenient form of notation introduced by Musiela [61] in 1993, where given a forward contract $f(t,T)$ with delivery time T, we reparametrize $f(t,T)$ to a form where we express the contract as a function of the time to delivery. This is done by defining $x = T - t$, and we get a forward contract on the form $f(t, T - t)$ which we will denote by $g(t,x)$. By doing this we will also motivate the forward dynamics (59) in chapter 8.4 and onwards. We will explain this by doing two examples.
Let us now consider a very general form of the forward price. We define

$$\mathrm{d}f(t,T) = \mu(t,T)\,\mathrm{d}t + \sigma(t,T)\,\mathrm{d}B(t), \tag{42}$$

for some suitably nice functions μ and σ. We can then reparameterize the forward curve by the time to maturity. That is, $x = T - t$ and define $g(t,x) := f(t,t+x)$. Then the domain of g becomes the (possibly unbounded) interval $[0,x]$.This very convenient notation was introduced by Musiela [61] in 1993. We start with an example where the spot price at time t, $S(t)$, follows an Ornstein-Uhlenbeck process denoted $X(t)$ at time t.

Example 7.2. Consider an Ornstein-Uhlenbeck process $t \mapsto X(t)$ given by

$$\mathrm{d}X(t) = \alpha\,(\mu - X(t))\,\mathrm{d}t + \sigma\,\mathrm{d}W(t),$$

with initial condition $X(0) = x_0$, $\alpha, \sigma \in \mathbb{R}_+$, $\mu \in \mathbb{R}$ and $t \mapsto W(t)$ is a Brownian motion under some equivalent martingale measure Q. By Itô's formula and variation of constants (see for example chapter 6.1), we obtain the solution

$$X(T) = e^{-\alpha(T-t)}X(t) + \mu\left(1 - e^{-\alpha(T-t)}\right) + \int_t^T e^{-\alpha(T-s)}\sigma\,\mathrm{d}W(s).$$

By appealing to no arbitrage arguments, we find that the price on a forward contract whose underlying follows X (that is, $S(t) = X(t)$ for all $t \geq 0$) at time t with delivery at time $T \geq t$ is

$$\begin{aligned} f(t,T) = &\ \mathbb{E}_Q[S(T)|\mathcal{F}_t] \\ &= e^{\alpha(T-t)}S(t) + \mu\left(1 - e^{-\alpha(T-t)}\right), \end{aligned}$$

since $S(t)$ is $\mathcal{F}_t$-measurable and the stochastic integral $\int_t^T e^{-\alpha(T-s)}\sigma\,\mathrm{d}W(s)$ is, due to the independent increment property of W, independent of $\mathcal{F}_t$ and therefore has expectation zero.

We will now introduce the Musiela parametrization, and defien $x = T - t$. Then

$$g(t,x) := f(t, t+x) = e^{-\alpha x} S(t) + \mu \left(1 - e^{-\alpha x}\right).$$

Itô's formula yields the differential dynamics of g as

$$\mathrm{d}g(t,x) = e^{-\alpha x} \alpha \left(\mu - S(t)\right) \,\mathrm{d}t + e^{-\alpha x} \sigma \,\mathrm{d}W(t).$$

Moreover, the derivative of g with respect to x is found as

$$\frac{\partial}{\partial x} g(t,x) = -\alpha e^{-\alpha x} S(t) + \alpha \mu e^{-\alpha x},$$

from which we deduce that

$$\mathrm{d}g(t,x) = \frac{\partial}{\partial x} g(t,x) \,\mathrm{d}t + e^{-\alpha x} \sigma \,\mathrm{d}W(t). \tag{43}$$

We can then define $\tilde{\sigma}(x) = e^{-\alpha x}\sigma$, and find

$$\mathrm{d}g(t,x) = \frac{\partial}{\partial x} g(t,x) \,\mathrm{d}t + \tilde{\sigma}(x) \,\mathrm{d}W(t), \tag{44}$$

which is known as the *Musiela stochastic partial differential equation*. As we will see, (44) is quite similar to the equations governing the forward dynamics of the elements g in the forward curve space H_w, which will be defined later. □

In the next example, we will consider a geometric model. This example is from the theory on interest rates.

Example 7.3. Suppose that the short rate $t \mapsto r(t)$ follows an Ornstein-Uhlenbeck process (Vasicek model). That is

$$\mathrm{d}r(t) = \left(b + \beta r(t)\right) \,\mathrm{d}t + \sigma \,\mathrm{d}W(t),$$

where $b \in \mathbb{R}$, $\beta \in \mathbb{R}_-$[12], $\sigma \in \mathbb{R}_+$ and $t \mapsto W(t)$ is Brownian motion under an equivalent martingale measure Q. It is known that the price of a zero coupon bond, $P(t,T)$, is given as

$$P(t,T) = \mathbb{E}_Q \left[\exp\left(-\int_t^T r(s) \,\mathrm{d}s \right) \right],$$

see for example Filipović [33]. Moreover, we also know that the long rates, denoted $f(t,T) = -\frac{\partial}{\partial T} \log(P(t,T))$. Using this, we may do similar calculations as

[12]This is a somewhat unusual way to define an Ornstein-Uhlenbeck process, but it is done this way in order to make the notation consistent with Filipović 's book [33].

above[13] and find that under the Musiela parametrization $g(t,x) = f(t,t+x)$ we get that

$$\mathrm{d}g(t,x) = \left(\frac{\partial}{\partial x} g(t,x) + \frac{\sigma^2}{\beta} 2^{2\beta x} - \frac{\sigma^2}{\beta} e^{\beta x}\right) \mathrm{d}t + \sigma e^{\beta x}\, \mathrm{d}W(t).$$

However, if do as in the previous example and define $\tilde{\sigma}(x) = \sigma e^{\beta x}$ we find that

$$\mathrm{d}g(t,x) = \left(\frac{\partial}{\partial x} g(t,x) + \tilde{\sigma}(x) \int_0^x \tilde{\sigma}(u)\, \mathrm{d}u\right) \mathrm{d}t + \tilde{\sigma}(x)\, \mathrm{d}W(t). \tag{45}$$

Moreover, in (45) we see that the part that differs from (43) is $\tilde{\sigma}(x) \int_0^x \tilde{\sigma}(u)\, \mathrm{d}u$, which we recognize as the no-arbitrage drift condition in the Heath, Jarrow & Morton framework! See for example **Theorem 6.1** in Filipović [33]. We may therefore re-express (45) in the same manner as in (44) and get

$$\mathrm{d}g(t) = \left(\frac{\partial}{\partial x} g(t,x) + \tilde{\mu}(x)\right) \mathrm{d}t + \tilde{\sigma}(x)\, \mathrm{d}W(t).$$

(7.3) then becomes the bridge to the differential equation governing the dynamics of the forward curves in H_w in the geometric case, in the sense that (7.3) motivates (88). □

[13]This calculation is somewhat involved, although not very difficult. It is therefore omitted.

8 Infinite dimensional stochastic analysis

In this section, we will cover some of the basics on infinite dimensional stochastic analysis, including Hilbert space-valued Wiener processes and its associated covariance operators. We will give a thorough introduction, which is mostly self contained, and prove some results that is otherwise (to my knowledge) just stated in the literature. We begin with a thorough definition of the Filipović space, which is a separable Hilbert space and the space for which the underlying processes on which we want to price derivatives takes values in.

8.1 The Filipović space

As stated in the introduction, the Filipović space will be the Hilbert space most suitable for our applications. This is due to the fact that we need a Hilbert space for which the evaluation functional δ_x is a continuous linear functional from $\mathbb{R}_+$ into $\mathbb{R}$. We also needed that the differentiation operator $\frac{\partial}{\partial x}$ must be a C_0 semi-group generating operator such that the shift semigroup $S(\cdot)$ is C_0. In a more financial context, the Filipović space also makes all its elements "flatten out" in the long run. The financial interpretation is that if we for example want to buy one power contract with delivery in 100 years, and another with delivery in 105 years, the price should be (more or less) the same. Formally, this means that each function h in this space converges to a limit $h(\infty)$ as $x \to \infty$. This is a consequence of Hölder's inequality, and is shown in the proof of **Theorem 5.1.1** in Filipović [34].

We start by defining the Filipović space, which we denote by H_w.

Definition 8.1 (Filipović Space)**.** Let $w : \mathbb{R}_+ \to [1, \infty]$ be a continuous and increasing weight function such that $w(0) = 1$. The Filipović space H_w is the space of all absolutely continous functions $g : \mathbb{R}_+ \to \mathbb{R}$ such that

$$\int_{\mathbb{R}_+} w(x) g'(x)^2 \, \mathrm{d}x < \infty. \tag{46}$$

The inner product on H_w is

$$\langle f, g \rangle = f(0)g(0) + \int_0^\infty w(x) g'(x) f'(x) \, \mathrm{d}x, \tag{47}$$

where $f, g \in H_w$.

It is shown in the proof of **Theorem 5.1.1** in Filipović [34] that H_w is isometrically isomorphic to the separable Hilbert space$\mathbb{R} \times L^2(\mathbb{R}_+)$, which implies that H_w is a separable Hilbert space as well. We will now define the norm on H_w.

Definition 8.2. The norm of H_w is

$$\|g\|_w^2 = \langle g, g \rangle. \tag{48}$$

We will now prove that the inner product and norm given above in (47) and (8.2) actually are an inner product and a norm. We start with the inner product. We use the definition of norm and inner product as defined in Lindstrøm [52].

Lemma 8.3 ((†))**.** *The function* $\langle \cdot, \cdot \rangle$ *given by* (47) *is an inner product on* H_w.

Proof: Let $f, g, h \in H_w$. The functions f and g are real valued, hence $\langle f, g \rangle = \langle g, f \rangle$ trivially.

$$\begin{aligned}
\langle f + g, h \rangle &= (f(0) + g(0))\, h(0) + \int_0^\infty w(x) \left(f'(x) + g'(x)\right) h'(x) \,\mathrm{d}x \\
&= f(0)h(0) + g(0)h(0) + \int_0^\infty w(x) f'(x) h'(x) \,\mathrm{d}x + \int_0^\infty g'(x) h'(x) \,\mathrm{d}x \\
&= \langle f, h \rangle + \langle g, h \rangle.
\end{aligned}$$

$$\begin{aligned}
\langle \beta f, g \rangle &= \beta f(0) g(0) + \beta \int_0^\infty w(x) f'(x) g'(x) \,\mathrm{d}x \\
&= \beta \left(f(0) g(0) + \int_0^\infty w(x) f'(x) g'(x) \,\mathrm{d}x \right) \\
&= \beta \langle f, g \rangle.
\end{aligned}$$

$$\begin{aligned}
\langle f, f \rangle &= f(0)^2 + \int_0^\infty w(x) f'(x)^2 \,\mathrm{d}x \\
&\geq f(0)^2 + \int_0^\infty f'(x)^2 \,\mathrm{d}x \\
&\geq 0,
\end{aligned}$$

where we see it is clear that we get zero if and only if $f(x) = 0$ for all $x \in \mathbb{R}_+$. ■

We can now do the same for the norm.

Proposition 8.4. *The function* $\|\cdot\|$ *given by* (8.2) *is a norm on* H_w.

Proof: 1. Nonnegativity: Since $\|g\|_w = \langle g, g\rangle$ the nonnegativity property follows from Lemma 8.3.
2. We have that

$$\begin{aligned}\|\beta g\|_w^2 &= \langle \beta g, \beta g\rangle \\ &= |\beta|^2 \langle g, g\rangle \\ &= |\beta|^2 \|g\|_w^2.\end{aligned}$$

Taking the square root on both sides yields the desired result.

3.

$$\begin{aligned}\|f+g\|_w^2 &= \langle f+g, f+g\rangle \\ &= \langle f, f\rangle + 2\langle f, g\rangle + \langle g, g\rangle \\ &\leq \|f\|_w^2 + 2\|f\|_w\|g\|_w + \|g\|_w^2 \\ &= \left(\|f\|_w + \|g\|_w\right)^2,\end{aligned}$$

hence $\|f+g\|_w \leq \|f\|_w + \|g\|_w$, and we have proved that $\|\cdot\|$ is a norm on H_w. ■

We have thus proven that the Filipović space is a normed space.

We will now prove that the Filipović space is a *locally convex topological vector space*. It is stated in McDonald & Weiss [55] that any normed space is a locally convex topological vector space. However, they simply state it without proof, and I have been unable to locate a proof in the literature. We will therefore prove that any normed space (and therefore the Filipović space) is a locally convex topological space. This is done so that we may later define a *convex risk measure* on the Filipović space.

We begin by defining a topological vector space and proving that any normed space is a topological vector space.

Definition 8.5 (Topological vector space)**.** A topological vector space is a vector space X over a field K (usually $\mathbb{R}$ or $\mathbb{C}$) endowed with a topology such that the maps $(x, y) \mapsto x + y$ and $(\alpha, x) \mapsto \alpha x$ are continuous from $X \times X$ and $K \times X$ to X.

We can now prove a simple result proving that any normed space is a topological vector space.

Lemma 8.6. *Any normed space is a topological vector space.*

Proof: Let X be a normed space, and let $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}} \subset X$ be such that $x_n \to x$ and $y_n \to y$ and let α be a scalar in X. Then

$$\|\alpha(x+y) - \alpha(x_n + y_n)\| \leq |\alpha| \left(\|x - x_n\| + \|y - y_n\|\right) \to 0,$$

This shows that both addition and scalar multiplication are continuous maps in X, and therefore X is a topological vector space. ■

We can now define a locally convex topological vector space.

Definition 8.7 (Locally convex topological vector space)**.** A *locally convex topological vector space* is a topological vector space that has a base for the topology consisting of convex sets.

From these definitions, it is easy to prove that every norm induces a metric given by $d(x, y) = \|x - y\|$ see for example Jameson [45], result 2.8 p. 21.

From Munkres [60], we also have the following

Definition 8.8 (Topology induced by metric)**.** Let d be a metric on Ω, then the collection of ϵ-balls $B_d(x, \epsilon) = \{y : d(x, y) < \epsilon\}$, for $x \in \Omega$ and $\epsilon > 0$, is a basis for a topology on Ω, called the *metric topology* induced by d.

We can now prove that any normed space is a locally convex topological vector space.

Theorem 8.9. *Any normed space is a locally convex topological vector space.*

Proof: Let us define $d(x, y) = \|x - y\|$. Then d is a metric on X. We have from above that the set $B_d(x, \epsilon) = \{y : d(x, y) < \epsilon\}$, for $x \in \Omega$ and $\epsilon > 0$ forms a basis for the metric topology on X. It then follows from **Theorem 5.14** in Folland [39] (along with the fact that any norm is a seminorm) that any normed space is a locally convex topological space. ■

From this it follows that the Filipović space is a locally convex topological vector space. We end this section about the topological properties of the Filipović space, which, even though they are not used in this thesis later, they are interesting in their own right.

Proposition 8.10. *The Filipović space is regular, second countable and Hausdorff.*

Proof: The Filipović space is both metrizable (as is any metric space) and separable (see **Theorem 5.1.1** in Filipović [34] for a proof), so it follows from Urysohn's metrization theorem that the Filipović space is regular, second countable and Hausdorff. ■

And the following corollary

Corollary 8.11. *The Filipović space is normal.*

Proof: This is a known result from topology, and follows the fact that the Filipović space is second-countable and regular. See for example **Theorem 32.1** in Munkres [60]. ■

We could also have proven the corollary above by simply appealing to the fact that the Filipović space is metrizable, and therefore by **Theorem 32.2** in Munkres [60] normal, as is any normed space.

8.2 Hilbert space valued Wiener processes and covariance operators

This subsection will for the most part follow Da Prato & Zabczyk [65].

Definition 8.12 (Wiener Process)**.** A real valued stochastic process $\{W(t) : t \geq 0\}$ is called a *Wiener process* if

1. W has continuous trajectories and $W(0) = 0$ with probability 1
2. W has independent increments
3. $\mathcal{L}(W(t) - W(s)) = \mathcal{L}(W(t - s)), t \geq s \geq 0$
4. $\mathcal{L}(W(t)) = \mathcal{L}(-W(t)), t \geq 0$

where $\mathcal{L}(X)$ denotes the law of X.

They also state an equivalent definition

Definition 8.13. A real valued stochastic process $\{W(t), t \geq 0\}$ with continuous trajectories is called a Wiener process if it is Gaussian and there exists $\sigma \geq 0$ such that $\mathbb{E}[W(t)] = 0$ and $\mathbb{E}[W(t)W(s)] = \sigma \min(s, t)$.

If E is a linear topological space, for example a Hilbert space of functions, then if $W(t)$ is an E-valued stochastic process satisfying either of the definitions above, then we say that $W(t)$ is an E-valued stochastic process.

Definition 8.14 (Gaussian measure on $\mathbb{R}^n$)**.** Let $n \in \mathbb{N}$ and $\mathcal{B}(\mathbb{R}^n)$ denote the Borel σ-algebra on $\mathbb{R}^n$. Let $\lambda : \mathcal{B}(\mathbb{R}^n) \to [0, \infty]$ be the standard n-dimensional Lebesgue measure. Then the *Gaussian measure* γ_n with mean $\mu \in \mathbb{R}^n$ and variance σ^2 is defined by

$$\gamma_n(A) = \left(2\pi\sigma^2\right)^{-n/2} \int_A \exp\left(-\frac{1}{2\sigma^2}\|x-\mu\|^2_{\mathbb{R}^n}\right) \,\mathrm{d}\lambda(x). \tag{49}$$

If $\mu = 0$, then γ^n is called a *centered Gaussian measure*.

We can now define Gaussian measures on any separable Banach space.

Definition 8.15 (Gaussian measure on separable Banach space)**.** Let E be a separable Banach space. A probability measure μ on $(E, \mathcal{B}(E))$ is said to be a Gaussian measure on $(E, \mathcal{B}(E))$ if the law of any arbitrary linear functional $h \in E^*$ considered as a random variable on $(E, \mathcal{B}(E))$ is a Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. If the law on each $h \in E^*$ is zero mean, μ is called a *centered Gaussian measure*.

Remark 8.16. We note that the definition above also holds for any separable Hilbert space, since all Hilbert spaces are Banach spaces.

Moreover, if μ is a Gaussian measure on a Hilbert space H, then there exists an element $m \in H$ and a linear operator Q such that for all $h, h_1, h_2 \in H$ we get

$$\int_H \langle h, x\rangle \mu(\,\mathrm{d}x) = \langle m, h\rangle,$$

and

$$\int_H \langle h_1, x-m\rangle\langle h_2, x-m\rangle \mu(\,\mathrm{d}x) = \langle Qh_1, h_2\rangle,$$

where

$$\int_H \langle h_1, x-m\rangle\langle h_2, x-m\rangle \mu(\,\mathrm{d}x) = Cov(h_1, h_2). \tag{50}$$

We will now prove that this representation is equal to the more familiar representation of the covariance, i.e

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]. \tag{51}$$

Proposition 8.17. *Equations* (50) *and* (51) *are equivalent.*

Proof: Let H be a Hilbert space, and denote its dual by H^*. Let $h \in H$ and let $f \in H^*$. By the definition of H^*, f is a continuous linear functional on H, making the expectation well defined. As written in the preliminaries section, we define $\mathbb{E}[f] = \int_H f(x) P(\mathrm{d}x)$. Using our known representation of the covariance (51), we find that

$$Cov(f, g) = \int_H (f(x) - \mathbb{E}[f])\,(g(x) - \mathbb{E}[g])\, P(\,\mathrm{d}x). \tag{52}$$

Moreover, the underlying field is $\mathbb{R}$, so we get from Riesz representation theorem that H and H^* are isometrically isomorphic, and every functional $f \in H^*$ can be uniquely expressed as $f(y) = \langle y, x\rangle$ for all $y \in H$. We can represent the expected values as a single element $x' \in H$. This is called the Pettis integral representation of the measure P. Moreover, $E[f] = f(x') = \langle x', x\rangle$ by Riesz representation theorem. We substitute this into (52) and find

$$\begin{aligned} Cov(f, g) &= \int_H (f(x) - \mathbb{E}[f])\,(g(x) - \mathbb{E}[g])\, P(\,\mathrm{d}x) \\ &= \int_H (\langle h, x\rangle - \langle h, m\rangle)\left(\langle h', x\rangle - \langle h', m\rangle\right) P(\,\mathrm{d}x) \\ &= \int_H \langle h, x - m\rangle \langle h', x - m\rangle P(\,\mathrm{d}x), \end{aligned}$$

which is what we wanted to prove.
In the last equality we used the linearity of the inner product. ∎

Using the notation above, the vector m is called the *mean* of μ and its *covariance operator* is Q. By the nonnegativity of the inner product we see that Q is a nonnegative operator. We also have from Da Prato & Zabczyk [65] that μ has the characteristic functional

$$CF(\mu)(\lambda) = \int_H \exp(i\langle\lambda, x\rangle)\mu(\,\mathrm{d}x) = \exp(i\langle\lambda, m\rangle - \frac{1}{2}\langle Q\lambda, \lambda\rangle),$$

where $i = \sqrt{-1}$. It therefore follows that μ is completely characterized by m and Q, and is therefore denoted $N(m, Q)$.
We will now show some further properties of the covariance operator.

Lemma 8.18. *The covariance operator Q is symmetric.*

Proof: Since the covariance is a bilinear form from $H \times H$ into $\mathbb{R}$, we find that

$$\begin{aligned} \langle Qx, y\rangle &= \int_H \langle x, z\rangle\langle y, z\rangle \,\mathrm{d}P(z) \\ &= \int_H \langle z, x\rangle\langle z, y\rangle \,\mathrm{d}P(z) \\ &= \langle x, Qy\rangle. \end{aligned}$$

■

We also have the following result, due to Vakhania & Tarieladze [73], which is stated here for convenience.

Lemma 8.19. *Let H be an arbitrary Hilbert space, μ a strong second order measure on H. Then the covariance operator $Q : H \to H$ of μ is a nuclear operator.*

Hence we have shown that the covariance operator is positive, symmetric and nuclear. The next result, due to Bosq [15] proves that the converse is true as well.

Theorem 8.20. *Let H be a Hilbert space. An operator $Q : H \to H$ is a covariance operator if and only if it is symmetric, positive and nuclear.*

Proof: See **Theorem 1.7** in Bosq [15]. ■

Moreover, it is true that the covariance operator of any strong second-order measure is always compact. See e.g Chobanjan & Tarieladze [22] or Baker & McKeague [4] for a set of necessary and sufficient conditions for the covariance operator to be compact. Q is also of trace class, i.e $\mathrm{Tr} Q < \infty$ (see e.g Da Prato & Zabczyk **Proposition 2.16** [65]). Also, it follows from the Principle of uniform boundedness (see e.g McDonald & Weiss [55] **Theorem 12.2**) that Q is continuous. In fact, for a Banach space B any symmetric operator $T : B \to B^*$ is continuous. Using Cauchy-Schwarz we find that the norm of Q (and of any positive symmetric operator mapping B into B^*) is given by $\|Q\| = \sup_{x \in B} |\langle x, Qx\rangle|$. We note that if B is a Hilbert space, then the positivity criterion can be dropped to find the norm. The last two results and their proofs can be found in Vakhania, Tarieladze & Chobanyan [74][14].

Now, let H and U be to real separable Hilbert spaces, and let $Q \in L(U)$ be a nonnegative, symmetric operator of trace class. Then, since $\langle W(t), u\rangle$ is a real valued Wiener process for each $u \in U$, we find that $\mathscr{L}(W(t))$ is a mean zero Gaussian measure. Da Prato & Zabczyk state (without proof) that

$$\mathbb{E}[\langle W(t), u\rangle\langle W(s), u\rangle] = t \wedge s\, \mathbb{E}[\langle W(1), u\rangle^2] = t \wedge s\langle Qu, u\rangle,$$

or more generally

$$\mathbb{E}[\langle W(t), u\rangle\langle W(s), v\rangle] = t \wedge s\, \mathbb{E}[\langle W(1, u)\rangle\langle W(1), v\rangle] = t \wedge s\langle Qu, v\rangle. \tag{53}$$

This is not that hard to prove, so we will prove the general statement.

[14] Vakhania et. al also show in **Corollary 2** (p. 173) that if we are in a separable Hilbert space H, then the identity operator $I : H \to H$ is the covariance operator of any probability measure of weak order two. This is, however a digression and is included only since it seems interesting to know.

Proposition 8.21. *If U is a separable Hilbert space and $u, v \in U$ and W is an U-valued Wiener process with covariance operator Q, then*

$$\mathbb{E}[\langle W(t), u\rangle\langle W(s), v\rangle] = t \wedge s\ \mathbb{E}[\langle W(1, u)\rangle\langle W(1), v\rangle] = t \wedge s\langle Qu, v\rangle.$$

Proof: By **Proposition 4.3** in Da Prato & Zabczyk [65], we know that for a Hilbert valued Wiener process $W(t)$, we may write

$$W(t) = \sum_{n\geq 1}\sqrt{\lambda_n}\beta_n(t)e_n,$$

where $\{\lambda_n\}$ is a sequence of eigenvalues of Q, and $\beta_n(t)$ is a standard Brownian motion $\beta \sim N(0, t)$. We find that

$$\begin{aligned}
\mathbb{E}[\langle W(t), u\rangle\langle W(s), v\rangle] = {} & \mathbb{E}\left[\left\langle\sum_{n\geq 1}\sqrt{\lambda_n}\beta_n(t)e_n, u\right\rangle\left\langle\sum_{k\geq 1}\sqrt{\lambda_k}\beta_k(s)e_k, v\right\rangle\right] \\
&= \sum_{n,k\geq 1}\sqrt{\lambda_n\lambda_k}\langle e_n, u\rangle\langle e_k, v\rangle\ \mathbb{E}[\beta_n(t)\beta_k(s)] \\
&= \sum_{n\geq 1}\lambda_n\langle e_n, u\rangle\langle e_n, v\rangle \min(t, s) \\
&= \min(t, s)\sum_{n\geq 1}\langle e_n, u\rangle\langle\lambda_n e_n, v\rangle \\
&= \min(t, s)\sum_{n\geq 1}\langle e_n, u\rangle\langle Qe_n, v\rangle \\
&= \min(t, s)\sum_{n\geq 1}\langle Q\langle e_n, u\rangle, v\rangle \\
&= \min(t, s)\left\langle Q\sum_{n\geq 1}\langle u, e_n\rangle e_n, v\right\rangle \\
&= \min(t, s)\langle Qu, v\rangle,
\end{aligned}$$

from which (53) follows. In the calculations above we have used that $\mathbb{E}[B(t)B(s)] = \min(t, s)$ (see appendix), we have used Parseval's theorem meaning that $u = \sum_{n\geq 1}\langle u, e_n\rangle e_n$ (see for example Lindstrøm [52]), and that $Qe_n = \lambda e_n$, since we know from the spectral theorem for self-adjoint compact operators that there exists an orthonormal basis for H consisting of eigenvectors of Q, see for example Lax [49]. We may therefore choose e_n to be an eigenvector of Q. ■

Remark 8.22. If H is separable, then we have from **Corollary 7.36** in Rynne & Youngson [71] that the orthonormal basis described in the proof above has the form $\{e_n\}_{n=1} \cup \{z_n\}_{n=1}$, where $\{e_n\}$ is an orthonormal basis of $\mathrm{Im}\bar{T}$ and $\{z_n\}$ is an orthonormal basis of Ker T.

We can now define a Q-Wiener process.

Definition 8.23. Let U be a Hilbert space. A U-valued stochastic process $\{W(t), t \geq 0\}$ is called a *Q-Wiener process* if

1. $W(0) = 0$ a.s
2. W has independent increments
3. W has continuous trajectories
4. $\mathcal{L}(W(t) - W(s)) = N(0, (t-s)Q)$ for all $t \geq s \geq 0$.

We will now illustrate these definitions and results with an example on the separable Hilbert space ℓ^2 of square summable functions. We find the covariance operator and characterize the Wiener process on ℓ^2.

Example 8.24. Let ℓ^2 denote the Hilbert space of square summable functions.

We define a linear operator $Q : \ell^2 \to \ell^2$ by

$$Qe_n = \frac{1}{n^2} e_n,\ n \in \mathbb{N},$$

where $\{e_n\}$ form an orthonormal basis for ℓ^2.

We know from **Theorem 5.3.11** in Lindstrøm [52]) that for any x in some complete vector space V with orthonormal basis $\{e_n\}_n$ we have $x = \sum_{n=1}^{\infty} x_n e_n$. In this example $V = \ell^2$, meaning that the completeness is trivially satisfied. Moreover, the inner product $\langle \cdot, \cdot \rangle$ on l^2 is given by $\langle x_n, y_n \rangle = \sum_{n=1}^{\infty} x_n y_n$. Therefore, for each $x \in l^2$

$$\begin{aligned} Qx &= Q \sum_{n=1}^{\infty} x_n e_n \\ &= \sum_{n \geq 1} x_n Q e_n \\ &= \sum_{n \geq 1} \frac{x_n e_n}{n^2} \\ &= \left\{ \frac{x_n}{n^2}, n = 1, 2, ... \right\}. \end{aligned}$$

From the calculations above we find that $\mathrm{Tr}Q = \sum_{n\geq 1}\langle Qe_n, e_n\rangle = \sum_{n\geq 1} Qe_n = \sum_{n\geq 1}\frac{1}{n^2} = \frac{\pi^2}{6}$ proving that Q is of trace class.

We can now characterize a Wiener process on l^2. We have by **Proposition 4.3** in Da Prato & Zabczyk [65] that a Wiener process $W(t)$ has the expansion

$$W(t) = \sum_{n\geq 1}\sqrt{\lambda_n}\beta_n(t)e_n,$$

where $\beta_n(t) = \frac{\langle W(t), e_n\rangle}{\sqrt{\lambda_n}}$ is a standard Wiener process (i.e $\beta_n(t) \sim N(0,t)$) and $\{\lambda_n\}$ is a sequence of nonnegative real numbers such that $Qe_n = \lambda_n e_n$. We see from the way we defined Q that $\lambda_n = \frac{1}{n^2}$.

We find that the expectation of the increment $W(t) - W(s)$ is

$$\mathbb{E}[W(t) - W(s)] = \mathbb{E}\left[\sum_{n\geq 1}\frac{1}{n}(\beta_n(t) - \beta_n(s))e_n\right] = 0,$$

since $\beta_n(\cdot)$ is a standard Wiener process.

Also, the variance of the norm of the increment is

$$\begin{aligned}
\mathrm{Var}[\|W(t) - W(s)\|_2] &= \sum_{n\geq 1}\frac{1}{n^2}\mathrm{Var}\left[\beta_n(t) - \beta_n(s)\right] \\
&= \sum_{n\geq 1}\frac{1}{n^2}(t-s) \\
&= (t-s)\frac{\pi^2}{6} \\
&= (t-s)\mathrm{Tr}Q,
\end{aligned}$$

where we have used that $\sum_{n\geq 1} 1/n^2 = \pi^2/6$, and we find that $\mathcal{L}(W(t)-W(s)) = N(0,(t-s)Q)$. We see that Q is symmetric by earlier arguments, and since it is everywhere defined it is a closed operator which then implies that Q is bounded. This in turn implies that Q is self-adjoint. Since $(\ell^2)* = \ell^2$, we get that Q is is symmetric and positive for all elements in $(\ell^2)^*$, meaning that Q is a covariance operator by Baker & McKeague [4]. Thus, $\{W(t), t \geq 0\}$ is a Q-Wiener process with covariance operator Q. $\square$

8.3 Infinite dimensional stochastic equations

In this section we will introduce stochastic equations in infinite dimensions. We will introduce results on existence and uniqueness for stochastic evolution equations driven by both general Lévy processes, and the special case when the Lévy process is a Wiener process.

8.3.1 The general case: Equations driven by square integrable martingales

The following section follows chapter 9 in Peszat & Zabczyk [64] closely, and may therefore be skipped entirely by readers familiar with these concepts.

Consider the equation

$$\mathrm{d}X(t) = (AX(t) + F(X))\ \mathrm{d}t + \sigma\,(t, X(t))\ \mathrm{d}M(t), \tag{54}$$

where $\mu : [0,t] \times H \to H$, $\sigma : [0,t] \times H \to L^2(U,H)$, and $A : D(A) \to H$ is some operator generating a strongly continuous semigroup $(S(t), t \geq 0)$. H and U are separable Hilbert spaces.

If we compare (54) with the theory from chapter 6, we see that (54) is a stochastic differential equation of the Ornstein-Uhlenbeck type. However, as infinite dimensional spaces have quite a lot more structure than finite dimensional spaces, we need to state some conditions on (54) in order to ensure that it is well defined. Equations of this kind usually have three types of solutions,

1. Mild solutions
2. Weak solutions
3. Strong solutions

We will define what these solutions are, and state the conditions on (54) needed to ensure we have solutions of a given kind.

We begin with defining mild solutions.

Definition 8.25. Let $X(0)$ be a square integrable $\mathcal{F}_{t_0}$-measurable random variable on H. A predictable process $X : [t_0, \infty) \times \Omega \to H$ is called a *mild solution* of (54) starting at time t_0 from $X(0)$ if for all $T \geq t_0$

$$\sup_{t\in[t_0,T]} \mathbb{E}\left[|X(t)|^2\right] < \infty \tag{55}$$

and

$$X(t) = S(t)X(0) + \int_0^t S(t-u)\mu(u, X(u))\,\mathrm{d}u + \int_0^t S(t-u)\sigma(u, X(u))\,\mathrm{d}M(u), \tag{56}$$

where S is the shift semigroup, L is a Lévy process and μ, σ are some well-behaving functions. The solution (56) holds P-almost surely.

Under the same conditions, we can state the definition of a strong solution.

Definition 8.26 (Strong solution)**.** A solution of (54) is called a strong solution if

$$P\left(\int_{t_0}^{T} \|X(s)\|_H + \|AX(s)\|_H + \|F(s, X(s))\|_H \,\mathrm{d}s < \infty\right) = 1$$

and

$$P\left(\int_{t_0}^{T} \|G(s, X(s))\|^2_{L_{HS}(U,H)} \,\mathrm{d}s < \infty\right) = 1$$

and

$$X(t) = X(0) + \int_0^t AX(s) + F(s, X(s)) \,\mathrm{d}s + \int_0^t G(s, X(s)) \,\mathrm{d}M(s)$$

with probability 1.

Remark 8.27. The integrals in (56) are all well-defined. We refer to **Remark 9.6** in Peszat & Zabczyk [64] for a discussion on this.

Remark 8.28. There are cases where mild solutions exist, but strong solutions do not.

In order to define weak solutions and state results on the uniqueness of the solutions, we need to state some conditions. We see that these conditions are to some degree analogous the the linear growth- and Lipschitz conditions for finite dimensional stochastic differential equations (see for example **Theorem 5.2.1** in Øksendal [62]).

(E1) $D(F)$ is dense in H and there is an integrable function $a : (0, \infty) \to (0, \infty)$ such that $\int_0^T a(t) \,\mathrm{d}t < \infty$ for all $T < \infty$ and for all $t > 0$ and $x, y \in D(F)$ we have

$$|S(t)F(x)|_H \leq a(t)(1 + |x|_H)$$

and

$$|S(t)\left(F(y) - F(x)\right)|_H \leq a(t)|y - x|_H.$$

(E2) $D(G)$ is dense in H and there is an integrable function $b : (0, \infty) \to (0, \infty)$ such that $\int_0^T b(t) \,\mathrm{d}t < \infty$ for all $T < \infty$ and for all $t > 0$ and $x, y \in D(G)$ we have

$$\|S(t)F(x)\|_{L_{HS}(\mathcal{H},H)} \leq b(t)(1 + |x|_H)$$

and

$$\|S(t)\left(F(y) - F(x)\right)\|_{L_{HS}(\mathcal{H},H)} \leq b(t)|y - x|_H.$$

We can now define weak solutions.

Definition 8.29. Let $t_0 \geq 0$ and assume that (E1) and (E2) hold and $X(0)$ is a square integrable $\mathcal{F}_{t_0}$-measurable random variable on H. Then, a predictable H-valued process $(X(t), t \geq t_0)$ is a weak solution of (54) if it satisfies (55) and for all $f \in D(A^*)$ and $t \geq t_0$

$$\langle f, X(t)\rangle_H = \langle f, X(0)\rangle_H + \int_{t_0}^{t} \langle A^* f, X(u)\rangle_H \, \mathrm{d}u + \int_{t_0}^{t} \langle f, F(X(u))\rangle_H \, \mathrm{d}u + \int_{t_0}^{t} \langle G^*(X(u)) f, \, \mathrm{d}M(u)\rangle_{\mathcal{H}}. \tag{57}$$

We do now have sufficient information to state a result on the uniqueness of the solutions.

Theorem 8.30. *Assume that (E1) and (E2) hold. Then the following are true*

(i) For all $t_0 \geq 0$ and $\mathcal{F}_{t_0}$-measurable random variables $X(0) \in H$ there exists a unique (up to modification) solution $X(\cdot, t_0, X(0))$ of (54).

(ii) For all $0 \leq t_0 \leq T < \infty$ there exists a number $K < \infty$ such that for all $x, y \in H$

$$\sup_{t \in [t_0, T]} \mathbb{E}\left[|X(t, t_0, x_1) - X(t, t_0, x_2)|_H^2\right] \leq K|x_1 - x_2|_h^2.$$

(iii) For all $0 \leq t_0 \leq t$ and $x \in H$, $\mathcal{L}(X(t, t_0, x))$ is independent of the choice of probability space.

Moreover, if we replace condition (E2) with

(E3) $D(G)$ is dense in H and there is an integrable function $b : (0, \infty) \to (0, \infty)$ such that $\int_0^T b(t)\, \mathrm{d}t < \infty$ for all $T < \infty$ and for all $t > 0$ and $x, y \in D(G)$ we have

$$\|F(x)\|_{L_{HS}(\mathcal{H},H)} \leq b(t)(1 + |x|_H)$$

and

$$\| (F(y) - F(x)) \|_{L_{HS}(\mathcal{H},H)} \leq b(t)|y - x|_H.$$

then

(iv) $X(t, t_0, X(0))$ has a càdlàg version.

Proof: This is proven in **Theorem 9.29** in Peszat & Zabczyk [64]. ■

Finally, to end this section we state the following monumental result, which is **Theorem 9.15** in Peszat & Zabczyk.

Theorem 8.31. *Assume that (E1) and (E2) hold. Then X is a mild solution if and only if X is a weak solution.*

Remark 8.32. If we let M be a Lévy process, we may remove the condition on square integrability. We refer to section 9.7 in Peszat & Zabczyk for a discussion. In that case they also state in **Theorem 9.35** that the solution X of (54) is a Markov process. In order to remove the square integrability condition, we have to introduce a sequence $\{\tau_n\}_{n\geq 1}$ of stopping times, which stops the process when it explodes. This is, however beyond the scope of this thesis.

8.3.2 Equations driven by Wiener processes

As in the general case, we consider the same evolution equation

$$\mathrm{d}X(t) = (AX(t) + F(t, X(t)))\ \mathrm{d}t + G(t, X(t))\,\mathrm{d}\mathbb{B}(t). \tag{58}$$

As in the section above, $A : D(A) \to H$ is an operator generating a strongly continuous semigroup, $F : \mathbb{R}_+ \times H \to H$, $G : \mathbb{R}_+ \times H \to L_{HS}(U, H)$ and $t \mapsto \mathbb{B}(t)$ is a mean zero H-valued Q-Wiener process. The conditions on the parameters ensuring the existence and uniqueness of a solution of equation (58) are similar to the general case, but somewhat relaxed.

Theorem 8.33. *Let $x, y \in H$ and let $K \in \mathbb{R}$ be finite. If F and G satisfy the Lipschitz bound*

$$\|F(t, x) - F(t, y)\|_H + \|G(t, x) - G(t, y)\|_{L_{HS}(U,H)} \leq K\|x - y\|_H$$

then there exists a unique (up to indistinguishability) H-valued mild solution $(X(t), t \geq 0)$ to (58) *such that for all $T \geq 0$ and $p > 2$ we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X(t)\|^p\right] < C_p\left(1 + \|X(0)\|^p\right).$$

Proof: This is **Theorem 4.3** in Carmona & Tehranchi [17], and we refer to them for a proof. ■

8.4 Applications to infinite dimensional modelling of financial derivatives

In this section we will focus on a special case of (54), where $F = 0$, $A = \frac{\partial}{\partial x}$ and $t \mapsto \mathbb{L}$ is an H-valued Lévy process.

$$\mathrm{d}g(t) = \frac{\partial}{\partial x} g(t)\,\mathrm{d}t + \sigma(t, g(t))\,\mathrm{d}\mathbb{L}(t). \tag{59}$$

Since the differential operator generates a strongly continuous semigroup (this is proven in **Theorem 5.1.1** in Filipović [34], and clarified later in this chapter in Theorems 8.36 and 8.37), we see from (56) that the solution of (59) is

$$g(t) = S(t)g(0) + \int_0^t S(t-u)\,\sigma(u, g(u))\,\mathrm{d}\mathbb{L}(u), \tag{60}$$

where $(S(t), t \leq T)$ is the shift semigroup, defined as $S(t)g(x) = g(x+t)$.

We will from now on assume that (59) is the dynamics governing the forward price, denoted as f, or more explicitly $f(t, T)$ where T denotes the delivery time. This equation will play a more vital role in subsequent chapters. In the remainder of this chapter we will introduce the model for a *forward contract* delivering some commodity over a time period $[\tau_1, \tau_2]$, which is defined as an integral where the integrand is a scaled version of the forward price f. We will then prove some technical results on the shift semigroup, and other needed results to begin the process of actually pricing such contracts, which will be done in the next chapter.

We know from Benth, Benth & Koekebakker [10] that the price of a forward contract, F, delivering some commodity over a time period $[T_1, T_2]$ may be expressed as

$$F(t, T_1, T_2) = \int_{T_1}^{T_2} \tilde{w}(u, T_1, T_2) f(t, u)\,\mathrm{d}u, \tag{61}$$

where $f(t, T)$ is the forward price at time t, and $\tilde{w}$ is some weight function. We then introduce the Musiela parametrization (as done in chapter 7.2) of (61), and define $x = T_1 - t$, $\ell = T_2 - T_1$, and set $g(t, y) = f(t, t+y)$. We then define (using the notation from Benth & Krühner [11])

$$G_\ell^w(t, x) = F(t; t+x, t+x+\ell) \tag{62}$$

$$= \int_x^{x+l} w_\ell(t, x, y) g(t, y)\,\mathrm{d}y, \tag{63}$$

for $w_\ell(t, x, y) = \tilde{w}(t+y; t+x, t+x+\ell)$, $x \leq y \leq x+\ell$, $x \geq 0$, $t \geq 0$.

For simplicity, we assume that w_ℓ is independent of t. Then

$$G^w_\ell(t,x) = \int_x^{x+\ell} w_\ell(y-x)g(t,y)\,\mathrm{d}y \tag{64}$$

$$=: \mathcal{D}^w_\ell\left(g(t,y)\right), \tag{65}$$

where

$$\mathcal{D}^w_\ell(g) = W_\ell(\ell)\mathrm{Id}(g) + \mathcal{I}^w_\ell(g), \tag{66}$$

for some arbitrary $g \in H_w$.

In (66), we have that the function $u \mapsto W_\ell(u)$ is given as

$$W_\ell(u) = \int_0^u w_\ell(x)\,\mathrm{d}x, \tag{67}$$

and that Id is the identity on H_w and that $\mathcal{I}^w_\ell(g)$ is a bounded linear integral operator on H_w, given by

$$\mathcal{I}^w_\ell(g) = \int_0^\infty q^w_\ell(\cdot,y)g'(y)\,\mathrm{d}y. \tag{68}$$

In (68) q^w_ℓ is the kernel $q^w_\ell(x,y) = \big(W_\ell(\ell) - W_\ell(y-x)1_{[0,l]}(y-x)\big)$.

Since $\mathcal{D}^w_\ell$ is a sum of a scaled identity operator and a bounded linear operator, it follows that $\mathcal{D}^w_\ell$ is a bounded linear operator on H_w as well.

We will now introduce the concept of a *weak derivative*, since the derivatives in the definition of the Filipović space is defined for weak derivatives.

Lemma 8.34. *If h has a weak derivative h', then there exists an absolutely continuous representation of h, still denoted by h such that*

$$h(x) - h(y) = \int_y^x h'(s)\,\mathrm{d}s. \tag{69}$$

Proof: This is from equation **(5.1)** in Filipović [34]. ■

In order to prove that the shift semigroup is strongly continuous, we need the following technical result. Its statement (without proof) can be found in Filipović [34], but for completion we will prove it.

Lemma 8.35. *For all $h \in H_w$ it holds true that*

$$h(x+t) - h(x) = t\int_0^1 h'(x+st)\,\mathrm{d}s. \tag{70}$$

Proof: We have from the preceeding lemma that $h(x+t)-h(x) = \int_x^{x+t} h'(u)\mathrm{d}u$. Then use substitution: $s = \frac{u-x}{t}$. Then we get that $\mathrm{d}u = t\mathrm{d}s$, and get $\int_x^{x+t} h'(u)\mathrm{d}u = t \int_0^1 h'(x + st)\, \mathrm{d}s$, which is what we wanted to prove. ■

We can now prove that the shift operator generates a strongly continuous semigroup on the Filipović space H_w. This statement is from Filipović , but the proof is rather short, so we have included a more detailed proof.

Theorem 8.36. *The shift operator* $\{S(t) : t \geq 0\}$ *given by* $S(t)g(x) = g(t + x)$ *is a strongly continuous semigroup on* H_w.

Proof: This proof is based on **Theorem 5.1.1** in [34], but details have been added by me.
First, we have to prove that $S(0)$ equals the identity on H_w. We find that

$$\begin{aligned} S(0)g(x) &= g(x) \\ &= \mathrm{Id}(g(x)), \end{aligned}$$

from which it follows that $S(0) = I$.

Second, we have to check that $S(t + s) = S(t)S(s)$ for all $t, s \geq 0$. By definition we find that

$$\begin{aligned} S(t + s)g(x) &= g(t + s + x) \\ &= S(t)g(s + x) \\ &= S(t)S(s)g(x), \end{aligned}$$

from which the result follows.

The last condition we need to prove, is that for any $g \in H_w$, it holds true that $\|S(t)g - g\| \to 0$ as $t \downarrow 0$. Using Lemma 8.35 and the Cauchy-Schwarz[15] inequality, we find that

[15] Cauchy-Schwarz-Bunyakovsky.

$$
\begin{aligned}
\|S(t)g - g\|_w^2 &= |g(t) - g(0)|^2 + \int_{\mathbb{R}_+} |g'(x+t) - g(x)|^2 w(x)\,\mathrm{d}x \\
&\overset{(70)}{=} |g(t) - g(0)|^2 + \int_{\mathbb{R}_+} |t \int_0^1 g''(x+st)\,\mathrm{d}s|^2 w(x)\,\mathrm{d}x \\
&\overset{(3)}{\leq} |g(t) - g(0)|^2 + t^2 \int_{\mathbb{R}_+} \int_0^1 1^2\,\mathrm{d}s \int_0^1 |g''(x+st)|^2\,\mathrm{d}s w(x)\,\mathrm{d}x \\
&= |g(t) - g(0)|^2 + t^2 \int_{\mathbb{R}_+} \int_0^1 |g''(x+st)|^2\,\mathrm{d}s w(x)\,\mathrm{d}x.
\end{aligned}
$$

We introduced the second derivative of g. This second derivative exists and is well defined since the set of all twice continuously differentiable whose first derivative has compact support is dense in H_w. Moreover, the absolute continuity of the Filipović space implies that any g H_w is measurable. We may therefore apply the Fubini-Tonelli theorem and find that

$$
\begin{aligned}
&|g(t) - g(0)|^2 + t^2 \int_{\mathbb{R}_+} \int_0^1 |g''(x+st)|^2\,\mathrm{d}s w(x)\,\mathrm{d}x \\
&= |g(t) - g(0)|^2 + t^2 \int_0^1 \int_{\mathbb{R}_+} |g''(x+st)|^2 w(x)\,\mathrm{d}x\,\mathrm{d}s \\
&= |g(t) - g(0)|^2 + t^2 \int_0^1 \int_{\mathbb{R}_+} |S(st)g''(x)|^2 w(x)\,\mathrm{d}x\,\mathrm{d}s \\
&\leq |g(t) - g(0)|^2 + t^2 |g'(st)|^2 \\
&\qquad + t^2 \int_0^1 \int_{\mathbb{R}_+} |S(st)g''(x)|^2 w(x)\,\mathrm{d}x\,\mathrm{d}s \\
&= |g(t) - g(0)|^2 + t^2 \int_0^1 \|S(st)g'\|_w^2\,\mathrm{d}s \\
&= |g(t) - g(0)|^2 + t^2 \|S(st)g'\|_w^2 \to 0 \text{ as } t \to 0,
\end{aligned}
$$

and it follows that $\|S(t)g - g\| \to 0$ as $t \downarrow 0$, which concludes the proof. ■

We end this section with a result that proves that $A = \frac{\partial}{\partial x}$ is the generator of the shift semigroup. This result is from Filipović [34] as well, but with added details by me.

Theorem 8.37 (Clearer version of **Corollary 5.1.1** in Filipović [34])**.** *The aforementioned semigroup has generator $A = \frac{\partial}{\partial x}$ and $D(A) = \{g \in H_w : g' \in H_w\}$.*

Proof: We want to show that $\|\frac{S(t)g-g}{t} - g'\| \to 0$ as $t \to 0$. We note that the construction of $D(A)$ implies the existence of $g"$. We find, using Lemma 8.35, the Cauchy-Schwarz-Bunyakovsky inequality and Fubini-Tonelli's theorem that

$$\begin{aligned}
&\|\frac{S(t)g-g}{t} - g'\|_w^2 = |\frac{g(t)-g(0)}{t} - g'(0)|^2 \\
&\quad + \int_{\mathbb{R}_+} |\frac{g'(t+x)-g'(x)}{t} - g"(x)|^2 w(x)\,\mathrm{d}x \\
&\overset{(70)}{=} \int_{\mathbb{R}_+} |\int_0^1 g"(x+st) - g"(x)\,\mathrm{d}s|^2 w(x)\,\mathrm{d}x \\
&\overset{(3)}{\leq} |\frac{g(t)-g(0)}{t} - g'(0)|^2 \\
&\quad + \int_{\mathbb{R}_+} \int_0^1 1^2\,\mathrm{d}s \int_0^1 |g"(x+st) - g"(x)|^2\,\mathrm{d}s w(x)\,\mathrm{d}x \\
&\overset{(4)}{=} |\frac{g(t)-g(0)}{t} - g'(0)|^2 \\
&\quad + \int_0^1 \int_{\mathbb{R}_+} |g"(x+st) - g"(x)|^2 w(x)\,\mathrm{d}x\,\mathrm{d}s \\
&\leq |\frac{g(t)-g(0)}{t} - g'(0)|^2 + \int_0^1 \|S(st)g'(x) - g'(x)\|_w^2\,\mathrm{d}s.
\end{aligned}$$

Moreover, we see that

$$\lim_{t\to 0} |\frac{g(t)-g(0)}{t} - g'(0)|^2 = 0$$

and we find by the dominated convergence theorem and the semigroup property of S that $\lim_{t\to 0} \int_0^1 \|S(st)g'(x) - g'(x)\|_w^2 w(x)\,\mathrm{d}s \to 0$. ■

9 Pricing of infinite dimensional derivatives

This section will deal with the pricing of infinite dimensional representations of financial derivatives. We will begin with some technical results.

The first result is stated in Benth & Krühner [11], but without proof. The proof is done by me.

Lemma 9.1. *Let $g \in H_w$. Then the following bound hold:*

$$\|\mathcal{D}_\ell^w(g)\|_w \leq \left(W_\ell(\ell) + \sqrt{W_\ell^2(\ell)\left(2 + \int_0^\ell w^{-1}(y)\,\mathrm{d}y\right) + 2c^2\ell^2} \right) \|g\|_w, \tag{71}$$

where $\mathcal{D}_\ell^w$ is defined as in (66).

Proof: We recall that we have defined

$$\mathcal{D}_\ell^w(g) = W_\ell(\ell)\mathrm{Id}(g) + \mathcal{I}_\ell{}^w(g),$$

where $W_\ell(u) = \int_0^u w_\ell(y)\,\mathrm{d}y$ and $\mathcal{I}_\ell{}^w(g)$ is some operator. From **Proposition 2.1** in Benth and Krühner [12], we find that $\mathcal{I}_\ell{}^w$ is a bounded linear operator on H_w. Following the notation of said proof, we define $\xi(x) := \mathcal{I}_\ell^w(g)(x) = \int_0^\infty q_\ell^w(x,y)g'(y)\mathrm{d}y$, hence $\xi(x) = \int_x^{x+\ell}(W_\ell(\ell) - W_\ell(y-x))g'(y)\mathrm{d}y$. Further, Benth & Krühner show that the Filipović -norm of ξ is given as

$$\|\xi\|_w \leq \|g\|_w \sqrt{2c^2\ell^2 + 2W_\ell^2(\ell) + W_\ell^2(\ell)\int_0^\ell w^{-1}(y)\,\mathrm{d}y}.$$

And therefore:

$$\begin{aligned}
\|\mathcal{D}_\ell^w(g)\|_w &= \|W_\ell(\ell)\mathrm{Id}(g) + \xi\|_w \\
&\leq W_\ell(\ell)\|g\|_w + \|\xi\|_w \\
&\leq W_\ell(\ell)\|g\|_w + \|g\|_w\sqrt{2c^2\ell^2 + 2W_\ell^2(\ell) + W_\ell^2(\ell)\int_0^\ell w^{-1}(y)\,\mathrm{d}y} \\
&\leq \left(W_\ell(\ell) + \sqrt{W_\ell^2(\ell)\left(2 + \int_0^\ell w^{-1}(y)\,\mathrm{d}y\right) + 2c^2\ell^2} \right)\|g\|_w,
\end{aligned}$$

which is what we wanted to prove. We used the triangle inequality in the first inequality, and used the bound on the Filipović -norm of ξ in the second. ■

The next result is a mixture of **Lemma 4.2.1** Filipović [34], and **Lemma 3.1** in Benth & Krühner [12].

Lemma 9.2. *For all* $u \in \mathbb{R}_+$ *there exists a number* $k(u)$ *such that* $\|\delta_x\|_H \leq k(u)$ *for all* $0 \leq x \leq u$. *Moreover, if we let* $h_x : \mathbb{R}_+ \to \mathbb{R}$ *be defined by* $h_x(y) = 1 + \int_0^{\min(y,x)} w(u)^{-1} \, \mathrm{d}u$, *then* $\|\delta_x\|_{\mathrm{op}}^2 = h_x(x)$.

Proof: See Filipović [34], pp 58-59 and Benth & Krühner [12] and the references therein. ∎

The next lemma is a simplified version of **Lemma 3.2** in Benth and Krühner [12], and stated here for convenience.

Lemma 9.3. *If* $\int_0^\infty w^{-1}(x) \, \mathrm{d}x < \infty$ *then* $\|g\|_\infty := \sup_{x \in \mathbb{R}_+} |g(x)| \leq c\|g\|_w$, *where* $c := \sqrt{1 + \int_0^\infty w^{-1}(x) \, \mathrm{d}x}$.

The next theorem is **Proposition 3.1** in Benth & Krühner [11], and provides a link to the infinite dimensional swap prices. However, it seems that the proof has some inconsistencies, hence a new proof was needed.

Theorem 9.4. *Let* $p : \mathbb{R} \to \mathbb{R}$ *be a function of at most linear growth. Then*

$$p\left(F\left(t, T_1, T_2\right)\right) = \mathcal{P}_{T_2 - T_1}^w \left(T_1 - t, g(t)\right), \tag{72}$$

where $\mathcal{P}_\ell^w : \mathbb{R}_+ \times H_w \to \mathbb{R}$ *is a nonlinear functional defined as*

$$\mathcal{P}_\ell^w \left(x, g\right) = p \circ \delta_x \circ \mathcal{D}_\ell^w \left(g\right). \tag{73}$$

Moreover,

$$\|\mathcal{P}_\ell^w(x, g)\|_\infty \leq c_\ell \left(1 + \|g\|_w\right), \tag{74}$$

for some ℓ*-dependent constant* $c_\ell > 0$.

Proof: We know that $F(t, T_1, T_2) = \mathcal{D}_{T_2 - T_1}^w \left(g(t)\right) (T_1 - t)$. Moreover,

$$p \circ \delta_x \circ \mathcal{D}_\ell^w(g) = p\left(\mathcal{D}_\ell^w(g)(x)\right),$$

and thus

$$p\left(F\left(t, T_1, T_2\right)\right) = p\left(\mathcal{D}_{T_2 - T_1}^w \left(g(t)\right) (T_1 - t)\right),$$

proving (72).
We can now prove the bound:
Using Lemma 9.3 and the linear growth of p, we get:

$$
\begin{aligned}
\|\mathcal{P}_\ell{}^w(x,g)\|_\infty &= \sup_{x\in\mathbb{R}_+} |\mathcal{P}_\ell{}^w(x,g)| \\
&\leq c\|\mathcal{P}_\ell{}^w(x,g)\|_w \\
&= \left(1+\int_0^\infty w^{-1}(x)\,\mathrm{d}x\right)^{1/2} \|\mathcal{P}_\ell{}^w(x,g)\|_w \\
&\leq \kappa\|\mathcal{P}_\ell{}^w(x,g)\|_w,
\end{aligned}
$$

where $\kappa = \left(1+\int_0^\infty w^{-1}(x)\,\mathrm{d}x\right)^{1/2}$. Next, using the definition of $\mathcal{P}_\ell{}^w$, the linear growth of p and Lemma 9.2, we find that

$$
\begin{aligned}
\leq \kappa\|\mathcal{P}_\ell{}^w(x,g)\|_w &= \kappa\|p\circ\delta_x\circ\mathcal{D}_\ell^w(g)\|_w \\
&\leq \kappa K\left(1+\|\delta_x\mathcal{D}_\ell^w(g)\|_w\right) \\
&\leq \kappa K\left(1+h_x(x)\|\mathcal{D}_\ell^w(g)\|_w\right).
\end{aligned}
$$

In Lemma 9.1, we found that

$$
\|\mathcal{D}_\ell^w(g)\|_w \leq \left(W_\ell(\ell)+\sqrt{W_\ell^2(\ell)\left(2+\int_0^\ell w^{-1}(y)\,\mathrm{d}y\right)+2c^2\ell^2}\right)\|g\|_w.
$$

We then define for brevity

$$
f(\ell) := \left(W_\ell(\ell)+\sqrt{W_\ell^2(\ell)\left(2+\int_0^\ell w^{-1}(y)\,\mathrm{d}y\right)+2c^2\ell^2}\right)
$$

and find that

$$
\begin{aligned}
\kappa K\left(1+h_x(x)\|\mathcal{D}_\ell^w(g)\|_w\right) &\leq \kappa K\left(1+h_x(x)f(\ell)\|g\|_w\right) \\
&\leq c_\ell\left(1+\|g\|_w\right),
\end{aligned}
$$

where c_ℓ is some positive ℓ-dependent constant. We therefore have that

$$
\|\mathcal{P}_\ell{}^w(x,g)\|_\infty \leq c_\ell\left(1+\|g\|_w\right),
$$

which is what we wanted to prove. ■

9.1 The forward model

We can now start with the actual pricing. Starting with the forward dynamics as introduced in the previous chapter. However, we start with arithmetic dynamics driven by an H-valued Q-Wiener process $\mathbb{B}$.

Our forward price model is now given as

$$\mathrm{d}g(t) = \frac{\partial}{\partial x} g(t) + \sigma(t, g(t))\,\mathrm{d}\mathbb{B}(t) \tag{75}$$

By forcing that σ is Lipschitz as seen in chapter 8, (75) has a unique mild solution as given in **Theorem 4.3** in Carmona & Tehranchi [20].

We see that in (75) we have A equal the differential operator and $\alpha(t,x) = 0$. Therefore, by the results on stochastic differential equations in chapter 8.3, if the differential operator generates a strongly continuous semigroup, and σ satisfies the Lipschitz bound, then there exists a unique, mild solution g to (75). We have already shown in Theorem 8.36 and Theorem 8.37 that $A = \frac{\partial}{\partial x}$ generates a strongly continuous semigroup, which implies that we have a unique, mild solution of the equation.We note that the dynamics of the forward model is modelled without drift. This is done to ensure that the process $t \mapsto g(t)$ is a martingale, which it is if and only if the drift term is zero. We refer to **Proposition 2.5** in Barth & Benth [7] for a proof of this result.

The next two results are **Lemma 3.3** in Benth & Krühner [12], where I have written new proofs.

Lemma 9.5. *For any $x \geq 0$, it holds true that*

$$S_x \mathcal{D}_\ell^w = \mathcal{D}_\ell^w S_x. \tag{76}$$

Proof: Let g be a function in H_w. Then

$$\begin{aligned}
\mathcal{D}_\ell^w S(t) g(x) &= \mathcal{D}_\ell^w g(x+t) \\
&= \int_0^\ell w_\ell(v)\,\mathrm{d}v \mathrm{Id}\,(g(x+t)) + \int_0^\infty q_\ell^w(x+t,y) g'(y)\,\mathrm{d}y \\
&= \int_0^\ell w_\ell(v)\,\mathrm{d}v \mathrm{Id}\,(S(t)g(x)) + \int_0^\infty S(t) q_\ell^w(x,y) g'(y)\,\mathrm{d}y \\
&= S(t)\int_0^\ell w_\ell(v)\,\mathrm{d}v \mathrm{Id}\,(g(x)) + S(t)\int_0^\infty q_\ell^w(x,y) g'(y)\,\mathrm{d}y \\
&= S(t)\mathcal{D}_\ell^w\,(g)\,(x).
\end{aligned}$$

■

Using these results, we can prove the following representation of a forward contract. This result is from Benth & Krühner as well, but with a new proof.

Lemma 9.6. *The forward contract may be expressed as*

$$F(\tau, T_1, T_2) = \delta_{T_1-t}\mathcal{D}^w_{T_2-T_1} g(t) + \int_t^\tau \delta_{T_1-u}\mathcal{D}^w_{T_2-T_1}\sigma(u, g(u))\, \mathrm{d}\mathbb{B}(u). \quad (77)$$

Proof: We know from earlier that $F(\tau, T_1, T_2) = \delta_{T_1-\tau}\mathcal{D}^w_{T_2-T_1} g(\tau)$. Applying the solution of (75) given by (56) and the preceding lemma we get:

$$\begin{aligned}
F(\tau, T_1, T_2) &= \mathcal{D}^w_{T_2-T_1}(g(\tau))(T_1-\tau) \\
&= \delta_{T_1-\tau}\mathcal{D}^w_{T_2-T_1}(g(\tau)) \\
&= \delta_{T_1-\tau}\mathcal{D}^w_{T_2-T_1}(S_{\tau-t}g(t)) + \delta_{T_1-\tau}\mathcal{D}^w_{T_2-T_1}\int_t^\tau S_{\tau-u}\sigma(u, g(u))\, \mathrm{d}\mathbb{B}(u) \\
&= \delta_{T_1-\tau}\mathcal{D}^w_{T_2-T_1}(S_{\tau-t}g(t)) + \int_t^\tau \delta_{T_1-\tau}S_{\tau-u}\mathcal{D}^w_{T_2-T_1}\sigma(u, g(u))\, \mathrm{d}\mathbb{B}(u) \\
&= \delta_{T_1-t}\mathcal{D}^w_{T_2-T_1} g(t) + \int_t^\tau \delta_{T_1-u}\mathcal{D}^w_{T_2-T_1}\sigma(u, g(u))\, \mathrm{d}\mathbb{B}(u),
\end{aligned}$$

where we after the third equality have inserted the solution $g(\tau)$ of (75) and used the commutation property (76) in the last equality.

■

The next result is **Theorem 2.1** in Benth & Krühner [12], and is stated here for convenience. The aim of this theorem is to provide a method of "pulling down" a process from an infinite dimensional space down to a finite dimensional space.

Lemma 9.7. *Let $n \in \mathbb{N}$ and H, U be separable Hilbert spaces. Let W be a square integrable and mean zero U-valued Wiener process with covariance $\mathcal{Q}$, which is a positive operator with finite trace. Assume that dim ran$(\mathcal{Q}) > n$ and $\mathcal{Q}$ is positive definite. Let $\Psi \in \mathcal{L}^2_W(H) := \bigcap\{\mathcal{L}^2_{W,T}(H) : T > 0\}$ where $\mathcal{L}_{W,T}(H)$ is the set of predictable linear operators between U and H such that $\int_0^T \mathrm{Tr}(\Psi(s)\mathcal{Q}\Psi(s)^*)\, \mathrm{d}s < \infty$, let $\mathcal{T} \in L(H, \mathbb{R}^n)$, the space of linear operators between H and $\mathbb{R}^n$, and define*

$$X(t) = \mathcal{T}\left(\int_0^t \Psi(s)\, \mathrm{d}W(s)\right).$$

Then there is an 1-dimensional standard Brownian motion B such that

$$X(t) = \int_0^t \sigma(s)\, \mathrm{d}B(s),$$

where $\sigma(s) = (\mathcal{T}\Psi(s)\mathcal{Q}\Psi(s)^*\mathcal{T}^*)^{1/2} \in \mathcal{L}_B^2(\mathbb{R}^n)$.

Using Lemma 9.7, we can prove the next result, which is equation (3.15) in Benth & Krühner, but is stated without proof. We have therefore proven it here.

Lemma 9.8. *The price of a forward contract delivering some commodity over a time period* $[T_1, T_2]$ *is given as*

$$F(\tau, T_1, T_2) = \delta_{T_1-t}\mathcal{D}_\ell^w g(t) + \int_t^\tau \tilde{\sigma}(s)\, \mathrm{d}B(s), \tag{78}$$

where $t \mapsto B(t)$ *is a 1-dimensional standard Brownian motion and* $\tilde{\sigma}^2(s) = (\delta_{T_1-s}\mathcal{D}^w{}_{T_2-T_1}\sigma(s)\mathcal{Q}\sigma^*(s)(\delta_{T_1-s}\mathcal{D}^w{}_{T_2-T_1})^*)(1)$.

Proof: According to Lemma 9.6, we have that

$$F(\tau, T_1, T_2) = \delta_{T_1-t}\mathcal{D}_{T_2-T_1}^w g(t) + \int_t^\tau \delta_{T_1-u}\mathcal{D}_{T_2-T_1}^w \sigma(u, g(u))\, \mathrm{d}\mathbb{B}(u).$$

Then define $\mathcal{T}$ to be the identity, and $\Psi(u) = \delta_{T_1-u}\mathcal{D}_{T_2-T_1}^w \sigma(u, g(u))$, then we find that

$$\tilde{\sigma}^2(s) = (\delta_{T_1-s}\mathcal{D}^w{}_{T_2-T_1}\sigma(s)\mathcal{Q}\sigma^*(s)(\delta_{T_1-s}\mathcal{D}^w{}_{T_2-T_1})^*)(1),$$

and the result follows. ■

We will now confine ourselves to Gaussian noise processes and nonrandom volatilities. We define the notation $V(t, g(t)) = e^{-r(\tau-t)}\,\mathbb{E}[p(F(\tau, T_1, T_2))]$. The next result is **Proposition 3.7** in Benth & Krühner [11]. The proof is a slightly altered version of the proof in the article, where some more details has been added.

Theorem 9.9. *The price of a claim on F is given by:*

$$V(t, g(t)) = e^{-r(\tau-t)}\,\mathbb{E}[p(m(g) + \xi X)], \tag{79}$$

where $X \sim N(0,1)$, $m(g) = \delta_{T_1-t}\mathcal{D}_{\mathcal{T}_2-\mathcal{T}}^w g(t)$, $\quad t \le \tau \le T_1$
and $\xi^2 = \int_t^\tau \tilde{\sigma}^2(s)\,\mathrm{d}s$.

Proof: We found in Lemma 9.8 that

$$F\,(\tau, T_1, T_2) = \delta_{T_1-t}\mathcal{D}_\ell^w g(t) + \int_t^\tau \tilde{\sigma}(s)\,\mathrm{d}B(s).$$

Then define $m(g(t)) := \delta_{T_1-t}\mathcal{D}_\ell^w g(t)$. Since we have defined σ to be nonramdom, we find from Wiersema [75] that the stochastic Itô integral $\int_t^\tau \tilde{\sigma}(s)\,\mathrm{d}B(s)$ is normally distributed with mean zero. Moreover, from the Itô isometry we find that the variance is $\xi^2 = \mathrm{Var}[\int_t^\tau \tilde{\sigma}(s)\,\mathrm{d}B(s)] = \int_t^\tau \tilde{\sigma}^2(s)\,\mathrm{d}s$. Therefore, we find that

$$\mathbb{E}\,[F\,(\tau, T_1, T_2)] = \delta_{T_1-t}\mathcal{D}_\ell^w g(t) = m(g(t))$$

and

$$\mathrm{Var}\,[F\,(\tau, T_1, T_2)] = \int_t^\tau \tilde{\sigma}^2(s)\,\mathrm{d}s,$$

where $\tilde{\sigma}$ is defined as in Lemma 9.7.
We therefore find that

$$F\,(\tau, T_1, T_2) = \delta_{T_1-t}\mathcal{D}_\ell^w g(t) + \int_t^\tau \tilde{\sigma}(s)\,\mathrm{d}B(s) = m(g(t)) + \int_t^\tau \tilde{\sigma}^2(s)\,\mathrm{d}s. \quad (80)$$

in distribution.
The payoff function of the claim is defined to be p, hence we may apply p to equation (80), and find the price by applying the discount factor and taking the expectation. ■

As a corollary, we can state an explicit Black-Scholes type formula for call options in our framework.

Corollary 9.10 ((†))**.** *A call option written on F with strike price K and exercise time τ, is*

$$V\,(t, g(t)) = e^{-r(\tau-t)}\left(\xi\phi\left(\frac{m(g(t)) - K}{\xi}\right) + (m(g(t) - K))\,\Phi\left(\frac{m(g(t)) - K}{\xi}\right)\right). \quad (81)$$

Here, Φ denotes the standard normal cumulative density function, and ϕ its derivative.

Proof: Since we are dealing with a call option, $p(\cdot) = \max(0, \cdot)$. And therefore

$$V\,(t, g(t)) = e^{-r(\tau-t)}\,\mathbb{E}\,[\max(0, m(g(t)) + \xi X - K)]\,.$$

We have that $V(t, g(t)) = 0$ whenever $X < (K - m(g(t))/\xi$. Hence

$$V(t, g(t)) = e^{-r(\tau-t)} \left(\frac{1}{\sqrt{2\pi}} \int_{\frac{K-m(g(t))}{\xi}}^{\infty} (m(g(t)) - K + \xi x)\, e^{-x^2/2}\, \mathrm{d}x \right)$$

$$= e^{-r(\tau-t)} \left(\xi\phi \left(\frac{m(g(t)) - K}{\xi} \right) + (m(g(t) - K))\, \Phi \left(\frac{m(g(t)) - K}{\xi} \right) \right).$$

■

Likewise, we can find a similar result for put options.

Corollary 9.11 ((†)). *A put option written on F with strike price K and exercise time τ is*

$$e^{-r(\tau-t)} \left((K - m(g(t)))\Phi \left(1 - \frac{m(g(t)) - K}{\xi} \right) - \xi\phi \left(1 - \frac{m(g(t)) - K}{\xi} \right) \right). \tag{82}$$

Proof: The price is

$$P(t, g(t)) = e^{-r(\tau-t)}\, \mathbb{E}\left[(K - (m(g(t)) + \xi X))^+\right].$$

So using the notation from the proof of the call option case, we find that

$$\begin{aligned} P(t, g(t)) &= \frac{e^{-r(\tau-t)}}{\sqrt{2\pi}} \int_d^{\infty} (K - m(g(t)) - \xi x) e^{-x^2/2}\, \mathrm{d}x \\ &= e^{-r(\tau-t)} \left((K - m(g(t))\Phi \left(1 - \frac{m(g(t)) - K}{\xi} \right) \right) \\ &\quad - \xi\phi \left(1 - \frac{m(g(t)) - K}{\xi} \right), \end{aligned}$$

which is what we wanted to prove. ■

We can also prove a result stating the relationship between call options and put options, a *put-call parity* of the Black-Scholes type.

Proposition 9.12 ((†) Put-call parity). *We have that*

$$C(t, g(t)) - P(t, g(t)) = e^{-r(\tau-t)} (F(t, T_1, T_2) - K). \tag{83}$$

Proof: We know that $\max(0, x - K) = (x - K) + \max(0, K - x)$. Therefore we find that since $P(t, g(t)) = e^{-r(\tau-t)}\, \mathbb{E}[(K - F(\tau, T_1, T_2))|\mathcal{F}_t]$:

$$\begin{aligned} C(t, g(t)) &= e^{-r(\tau-t)}\, \mathbb{E}[(F(\tau, T_1, T_2) - K)^+|\mathcal{F}_t] \\ &= e^{-r(\tau-t)}\, \mathbb{E}[F(\tau, T_1, T_2) - K|\mathcal{F}_t] \\ &\quad + e^{-r(\tau-t)}\, \mathbb{E}[(K - F(\tau, T_1, T_2))^+|\mathcal{F}_t] \\ &= P(t, g(t)) - Ke^{-r(\tau-t)} + e^{-r(\tau-t)}\, \mathbb{E}[F(\tau, T_1, T_2)|\mathcal{F}_t], \end{aligned}$$

and therefore

$$C(t, g(t)) - P(t, g(t)) = e^{-r(\tau-t)} \left(F(t, T_1, T_2) - K\right),$$

which is what we wanted to prove. ■

In the next result, we will consider a calendar spread option written on two contracts with different delivery periods on the same underlying. That is, the delivery periods are $[T_1, T_2]$ and $[S_1, S_2]$, and the option pays

$$p\left(F(\tau, T_1, T_2), F(\tau, T_1, T_2)\right)$$

at the exercise time $\tau \leq \min(T_1, S_1)$. We assume that the volatility σ is continuous, deterministic and square-integrable.

Theorem 9.13 ((†) Pricing calendar spread options on power forwards)**.** *Let $F(\cdot, T_1, T_2)$ and $F(\cdot, S_1, S_2)$ be two forward contracts delivering the same commodity over two different time periods $[T_1, T_2]$ and $[S_1, S_2]$ respectively. Denote by $CS(\cdot, g)$ the price of a calendar spread option written on F. The price is then*

$$CS(t, g) = e^{-r(\tau-t)} \left(\Phi(d) \left(F(t, T_1, T_2) - F(t, S_1, S_2)\right) + \xi\varphi(d)\right), \qquad (84)$$

where $d = \frac{1}{\xi}(F(t, T_1, T_2) - F(t, S_1, S_2))$ and

$$\xi^2 = \int_0^\tau \left((\Psi(u, T) Q \Psi^*(u, T))^{1/2} - (\Psi(u, S) Q \Psi^*(u, S))^{1/2} \right)^2 \mathrm{d}u,$$

where $\Psi(u, T) = \delta_{T_1-t} \mathcal{D}^w_{T_2-T_1} \sigma(u)$.

Proof: Define $\Psi(u, T) := \delta_{T_1-t} \mathcal{D}^w_{T_2-T_1} \sigma(u)$. Then, using Lemma 9.6 we get the following representation for F:

$$F(\tau, T_1, T_1) = \delta_{T_1-t} \mathcal{D}^w_{T_2-T_1} g(t) + \int_t^\tau \Psi(u, T) \, \mathrm{d}\mathbb{B}(u) \qquad (85)$$

$$F(\tau, S_1, S_1) = \delta_{S_1-t} \mathcal{D}^w_{S_2-S_1} g(t) + \int_t^\tau \Psi(u, S) \, \mathrm{d}\mathbb{B}(u). \qquad (86)$$

Using Lemma 9.7 we find that we may represent (85) and (86) as

$$F(\tau, T_1, T_2) = \delta_{T_1-t} \mathcal{D}^w_{T_2-T_1} g(t) + \int_t^\tau \tilde{\sigma}(u, T) \, \mathrm{d}B(u)$$

$$F(\tau, S_1, S_2) = \delta_{S_1-t} \mathcal{D}^w_{S_2-S_1} g(t) + \int_t^\tau \tilde{\sigma}(u, S) \, \mathrm{d}B(u),$$

where B is a 1-dimensional Brownian motion and $\tilde{\sigma}^2(u,\cdot) = \Psi(u,\cdot)Q\Psi(u,\cdot)^*$.

Since the option is written on the spread of $F(\cdot, T_1, T_2)$ and $F(\cdot, S_1, S_2)$, the price is

$$\begin{aligned} CS(t,g) &= e^{-r(\tau-t)}\,\mathbb{E}\left[(F(\tau,T_1,T_2) - F(\tau,S_1,S_2))^+\,|\mathcal{F}_t\right] \\ &= e^{-r(\tau-t)}\,\mathbb{E}\left[(F(\tau,T_1,T_2) - F(\tau,S_1,S_2))^+\right], \end{aligned}$$

where we in the last equation used that $F(t,\cdot,\cdot)$ is $\mathcal{F}_t$-measurable and that $\int_t^\tau \tilde{\sigma}(u)\mathrm{d}B(u)$ is independent of $\mathcal{F}_t$.

From above, it is clear that $F(\tau,T_1,T_2)-F(\tau,S_1,S_2) = F(t,T_1,T_2)-F(t,S_1,S_2)+\int_t^\tau \tilde{\sigma}(u,T) - \tilde{\sigma}(u,S)\,\mathrm{d}B(u)$. Using the standard Itô isometry, we find that in distribution

$$\begin{aligned} F(\tau,T_1,T_2) - F(\tau,S_1,S_2) &\stackrel{\mathrm{d}}{=} F(t,T_1,T_2) - F(t,S_1,S_2) \\ &\quad + \left(\int_t^\tau (\tilde{\sigma}(u,T) - \tilde{\sigma}(u,S))^2\,\mathrm{d}u\right)^{1/2} Z, \end{aligned}$$

where $Z \sim N(0,1)$. If we for convenience define $m(t) = F(t,T_1,T_2)-F(t,S_1,S_2)$ and $\xi^2 = \int_t^\tau (\tilde{\sigma}(u,T) - \tilde{\sigma}(u,S))^2\,\mathrm{d}u$, we find that

$$F(\tau,T_1,T_2) - F(\tau,S_1,S_2) \stackrel{\mathrm{d}}{=} m(t) + \xi Z.$$

In order to get the expectation above to be non-negative, we must have that

$$Z \geq -m(t)/\xi =: -d.$$

Then we find that

$$\begin{aligned} CS(t,g) &= e^{-r(\tau-t)}\,\mathbb{E}[(m(t)+\xi Z)^+] \\ &= e^{-r(\tau-t)}\frac{1}{\sqrt{2\pi}}\int_{-d}^\infty (m(t)+\xi z)\,e^{-z^2/2}\,\mathrm{d}z \\ &= e^{-r(\tau-t)}\,(m(t)\Phi(d) + \varphi(d)\xi)\,, \end{aligned}$$

which is exactly (84).

■

Remark 9.14. The proof of Theorem 9.13 may also be done without using Lemma 9.7. Define $\Psi(u,T)$ as above, and note that since we have assumed that σ is continuous, $\Psi \in H_w^*$. We find that

$$F(\tau,T_1,T_2) - F(\tau,S_1,S_2) = m(t) + \int_t^\tau \Psi(u,T) - \Psi(u,S)\,\mathrm{d}\mathbb{B}(s).$$

In order to get an expression for $F(\tau, T_1, T_2) - F(\tau, S_1, S_2)$ in distribution, we must compute the standard deviation of $F(\tau, T_1, T_2) - F(\tau, S_1, S_2)$. We find that

$$\begin{aligned}\xi^2 &:= \operatorname{Var}\left[F(\tau, T_1, T_2) - F(\tau, S_1, S_2)\right] \\ &= \mathbb{E}\Bigg[\left(\int_t^\tau \Psi(u, T)\,\mathrm{d}\mathbb{B}(u)\right)^2 + \left(\int_t^\tau \Psi(u, S)\,\mathrm{d}\mathbb{B}(u)\right)^2 \\ &\qquad - 2\int_t^\tau \Psi(u, T)\,\mathrm{d}\mathbb{B}(u)\int_t^\tau \Psi(u, S)\,\mathrm{d}\mathbb{B}(u)\Bigg].\end{aligned}$$

Define $X_T := \int_t^\tau \Psi(u, T)\,\mathrm{d}\mathbb{B}(u)$ and $X_S := \int_t^\tau \Psi(u, S)\,\mathrm{d}\mathbb{B}(u)$. Then

$$\begin{aligned}\mathbb{E}[X_T X_S] &= \mathbb{E}\left[\lim_{|u_i|,|u_j|\to 0}\sum_{i,j}\Psi(u_i, T)\left(\mathbb{B}(u_{i+1}) - \mathbb{B}(u_i)\right)\Psi(u_j, S)\left(\mathbb{B}(u_{j+1}) - \mathbb{B}(u_j)\right)\right] \\ &\overset{\mathrm{DCT}}{=} \lim_{|u_i|,|u_i|}\mathbb{E}\left[\sum_{i,j}\Psi(u_i, T)\mathbb{B}(\Delta u_i)\Psi(u_j, S)\mathbb{B}(\Delta u_j)\right] \\ &\overset{(*)}{=} \lim_{|u_i|,|u_i|}\mathbb{E}\left[\sum_i \Psi(u_i, T)\mathbb{B}(\Delta u_i)\Psi(u_i)\mathbb{B}(\Delta u_i)\right] \\ &\overset{(**)}{=} \lim_{|u_i|\to 0}\mathbb{E}\left[\sum_i \langle f(u_i, T), \mathbb{B}(\Delta u_i)\rangle\langle f(u_i, S), \mathbb{B}(\Delta u_i)\rangle\right] \\ &\overset{(***)}{=} \lim_{|u_i|\to 0}\mathbb{E}\left[\sum_i \langle f(u_i, T), \mathbb{B}(1)\rangle\langle f(u_i, S), \mathbb{B}(1)\rangle\Delta u_i\right] \\ &\overset{(***)}{=} \lim_{|u_i|\to 0}\sum_i \langle Qf(u_i, T), f(u_i, S)\rangle\Delta u_i \\ &= \int_t^\tau \langle Qf(u, T), f(u, S)\rangle\,\mathrm{d}u \\ &\overset{(****)}{=} \int_t^\tau f(u, T)^* Qf(u, S)(1)\,\mathrm{d}u.\end{aligned}$$

In the calculations above, we applied the Dominated convergence theorem since $\Psi \in H_w^*$ (i.e it is continuous and therefore bounded), and we also used the stationarity property of the Wiener process to get that $\mathbb{B}(u_{i+1}) - \mathbb{B}(u_i) = \mathbb{B}(u_{i+1} - u_i) =: \mathbb{B}(\Delta u_i)$. In equality $(*)$ we used that the expectation is zero whenever $i \neq j$ due to the independent increment property of the Wiener process. In equality $(**)$ we applied the Riesz-Fréchet representation theorem, which we may do since $\Psi \in H_w^*$ and $\mathbb{B} \in H_w$. In equalities $(***)$ we applied Proposition 8.21. Finally, in equality $(****)$ we used the fact that the covariance operator Q is self-adjoint. From this it is clear that $\mathbb{E}[X_T^2] = \int_t^\tau f(u,T)^* Q f(u,T)(1)\,\mathrm{d}u$ and that $\mathbb{E}[X_S] = \int_t^\tau f(u,S)^* Q f(u,S)(1)\,\mathrm{d}u$. We have therefore found ξ in the distributional representation $F(\tau,T_1,T_2) - F(\tau,S_1,S_2) = m(t) + \xi Z$, and we may proceed with the calculations as above. (end of remark).

The results up to this point in this chapter are done for arithmetic equations where the noise is driven by a H_w-valued Q-Wiener process. We will now comment slightly on geoemetric equations. We know from Benth & Krühner [11] that the Filipović space is closed under exponentiation, meaning that if $f \in H_w$, then $\exp(f) := \sum_{k=0}^{\infty} \frac{f^k}{k!} \in H_w$. We may therefore define the following model for the forward prices:

$$g(t) = \exp(\tilde{g}(t)), \tag{87}$$

where

$$\mathrm{d}\tilde{g}(t) = \left(\frac{\partial \tilde{g}(t)}{\partial x} + \mu(t)\right)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}\mathbb{B}(t), \tag{88}$$

where μ is a predictable Bochner integrable stochastic process, added to ensure that the dynamics are arbitrage-free. Benth & Krühner prove that the dynamics are arbitrage-free if $\mu(t,x) = -\frac{1}{2}\delta_x \sigma(t) Q \sigma^*(t) \delta_x^*(1)$. Using this drift condition, they prove in **Lemma 3.14** that

$$g(T) = S(T-t)g(t) + \int_t^T S(T-s)\hat{\sigma}(s, f(s))\,\mathrm{d}\mathbb{B}(s),$$

where $\hat{g}(s,x)h(x) = g(x)\sigma(s)h(x)$ for $g,h \in H_w$ and $x(s) \geq 0$. They also state without proof that

$$F(T,T_1,T_2) = \delta_{T_1-t}\mathcal{D}_\ell^w g(t) + \int_t^T \delta_{T_1-s}\mathcal{D}_\ell^w \hat{\sigma}(s, f(s))\,\mathrm{d}\mathbb{B}(s).$$

We will therefore do a proof.

Lemma 9.15. *In the geometric case, the forward price may be expressed as*

$$F(T,T_1,T_2) = \delta_{T_1-t}\mathcal{D}_\ell^w g(t) + \int_t^T \delta_{T_1-s}\mathcal{D}_\ell^w \hat{\sigma}(s, f(s))\,\mathrm{d}\mathbb{B}(s).$$

Proof: We know from earlier that $F(\tau,T_1,T_2) = \delta_{T_1-\tau}\mathcal{D}^w{}_{T_2-T_1} g(\tau)$ in general.

This leads to

$$\begin{aligned}
F(\tau, T_1, T_2) &= \delta_{T_1-\tau}\mathcal{D}^{w}{}_{T_2-T_1}g(\tau) \\
&= \delta_{T_1-\tau}\mathcal{D}^{w}{}_{T_2-T_1}S(\tau-t)g(t) \\
&\quad + \delta_{T_1-\tau}\mathcal{D}^{w}{}_{T_2-T_1}\int_t^\tau S(\tau-u)\hat{\sigma}(u)\,\mathrm{d}\mathbb{B}(u) \\
&= \delta_{T_1-t}\mathcal{D}^{w}{}_{T_2-T_1}g(t) + \int_t^\tau \delta_{T_1-\tau}\mathcal{D}^{w}{}_{T_2-T_1}\hat{\sigma}(u)\,\mathrm{d}\mathbb{B}(u),
\end{aligned}$$

which is what we wanted to prove. In the last equality we have used the commutation property from (76). ■

We may then apply Lemma 9.8 and write

$$F(\tau, T_1, T_2) = \delta_{T_1-t}\mathcal{D}^{w}{}_{T_2-T_1}g(t) + \int_t^\tau \tilde{\sigma}(s)\,\mathrm{d}B(s),$$

where $\tilde{\sigma}$ is defined in the same way as in Lemma 9.8.

This leads to the following results on option pricing (both results are stated in Benth & Krühner [11], but neither result is proved), which hold for non-random volatilities. These results hold for assets with fixed delivery time. Note that the statement in Benth & Krühner is wrong.[16] Their statement about the drift term has wrong sign at is multiplied with $\frac{1}{2}$ meaning that their statement about $\hat{m}$ should be

$$\hat{m}(g) = \tilde{g}(T-t) + \int_t^\tau \mu(s)(T-s)\,\mathrm{d}s.$$

Likewise, the term ξ is stated as if the contract has a delivery period and not fixed delivery. We will prove the statement in a similar, but somewhat different form. In the process, we also manage to prove the drift condition for our case, in a very different manner than what is proven in **Lemma 3.13** in Benth & Krühner [11].

Theorem 9.16. *Let $f(t,T)$ be a forward contract delivering some commodity at time T. Then we have the following representation*

$$\begin{aligned}
f(t,T) = f(0,T)\exp\bigg(&-\frac{1}{2}\int_0^t \delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*\,\mathrm{d}s \\
&+ \int_0^t (\delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*)^{\frac{1}{2}}\,\mathrm{d}B(s)\bigg).
\end{aligned} \tag{89}$$

[16]It would probably be more fair to call this an oversight instead of an error.

Moreover, a European option on $f(t,T)$ then has price

$$V(t,g) = e^{-rT}\,\mathbb{E}\left[p\left(f(0,T)\exp\left(-\xi^2 + \xi Z\right)\right)\right],$$

where $Z \sim N(0,1)$ and $\xi^2 = \int_0^t \delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*\,\mathrm{d}s$.

Proof: We know that the solution of (88) is given as the predictable mild solution

$$\tilde{g}(t) = S(t)g_0 + \int_0^t S(t-s)\mu(s)\,\mathrm{d}s + \int_0^t S(t-s)\sigma(s)\,\mathrm{d}\mathbb{B}(s).$$

Let us now consider a forward contract $f(t,T)$ with delivery time T. Since we are dealing with a geometric model, we have that

$$f(t,T) = \exp\left(\delta_{T-t}\tilde{g}(t)\right).$$

And therefore

$$\begin{aligned} f(t,T) &= \exp\Bigg(\delta_{T-t}S(t)g(0) + \int_0^t \delta_{T-t}S(t-s)\mu(s)\,\mathrm{d}s \\ &\qquad + \int_0^t \delta_{T-t}S(t-s)\sigma(s)\,\mathrm{d}\mathbb{B}(\mathbb{s})\Bigg) \\ &= f(0,T)\exp\left(\int_0^t \delta_{T-s}\mu(s)\,\mathrm{d}s + \int_0^t \delta_{T-s}\sigma(s)\,\mathrm{d}\mathbb{B}(s)\right). \end{aligned} \tag{90}$$

We can now apply **Theorem 2.1** in Benth & Krühner [12] to (90), and find that

$$f(t,T) = f(0,T)\exp\left(\int_0^t \delta_{T-s}\mu(s)\mathrm{d}s + \int_0^t (\delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*)^{\frac{1}{2}}\,\mathrm{d}B(s),\right.$$

where B is a 1-dimensional Brownian motion.

In order to have well defined option prices, the mapping $t \mapsto f(t,T)$ must be a martingale. To get this condition, we must force the drift μ to such that the exponential part of $f(t,T)$ has expectation 1. Since the exponent is Gaussian, we see that we need to have

$$\int_0^t \delta_{T-s}\mu(s)\,\mathrm{d}s + \frac{1}{2}\int_0^t \delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*\,\mathrm{d}s,$$

which in turn shows that the drift term must be

$$-\frac{1}{2}\int_0^t \delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*\,\mathrm{d}s, \tag{91}$$

proving **Lemma 3.13** in Benth & Krühner [11].
The forward price has therefore explicit representation

$$f(t,T) = f(0,T)\exp\bigg(-\frac{1}{2}\int_0^t \delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^* \,\mathrm{d}s \\ + \int_0^t (\delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^*)^{\frac{1}{2}} \,\mathrm{d}B(s)\bigg). \tag{92}$$

In distribution, we may represent (92) as

$$f(t,T) = f(0,T)\exp\left(-\frac{1}{2}\xi^2 + \xi Z\right),$$

where $Z \sim N(0,1)$ and

$$\xi^2 = \int_0^t \delta_{T-s}\sigma(s)Q(\delta_{T-s}\sigma(s))^* \,\mathrm{d}s,$$

proving the first assertion.

We may now apply standard option prices techniques, and a payoff function p to get the desired result, which is that the price of an option on the forward price with payoff p at time zero is represented as

$$V(0,g) = e^{-rT}\,\mathbb{E}\left[p\left(f(0,T)\exp\left(-\frac{1}{2}\xi^2 + \xi Z\right)\right)\right],$$

which is what we wanted to prove.

■

We can then prove the following corollary on the case when the payoff function p is the payoff function of call option. This is also a result from Benth & Krühner stated without proof.

Corollary 9.17. *If $p = \max(0, x-k)$, then we recover the Black-76 formula.*

Proof: Define $f_0 := f(0,T)$. Since the payoff function $p(x) = (x-K)^+$, where K is the strike price, the quantity we need to compute is

$$C := \mathbb{E}\left[(f(t,T) - K)^+\right].$$

Using the distribution representation of f as found in the proof of Theorem 9.16, we see that we must have

$$f_0 \exp\left(-\frac{1}{2}\xi^2 + \xi Z\right) - K > 0,$$

in order to have a nonzero expectation.
Consequently, we find that

$$Z > \frac{\log\left(\frac{K}{f_0}\right) + \frac{1}{2}\xi^2}{\xi} =: -d.$$

Using standard option pricing arguments, we find that the price is

$$\begin{aligned}
C &= \int_{-d}^{\infty} \left(f_0 e^{\xi^2/2 + \xi z} - K \right) \varphi(z)\, \mathrm{d}z \\
&= -K\Phi(d) + \frac{f_0}{\sqrt{2\pi}} \int_{-d}^{\infty} e^{-\frac{1}{2}(z^2 - 2\xi z + \xi^2)}\, \mathrm{d}z \\
&= -K\Phi(d) + \frac{f_0}{\sqrt{2\pi}} \int_{-d}^{\infty} e^{-\frac{1}{2}(z-\xi)^2}\, \mathrm{d}z \\
&= -K\Phi(d) + \frac{f_0}{\sqrt{2\pi}} \int_{-(d+\xi)}^{\infty} e^{-u^2/2}\, \mathrm{d}u \\
&= f_0\Phi(d+\xi) - K\Phi(d).
\end{aligned}$$

We may then multiply with the discount factor and conclude that the price of a call option of $f(t,T)$ at time 0 with strike price K is

$$e^{-rT}\left(f_0\Phi(d+\xi) - K\Phi(d) \right),$$

which we recognize as the Black-76 formula. ■

Like in the arithmetic case, we may derive a put call parity relationship:

Proposition 9.18 (†). *We have that*

$$C(t,g) - P(t,g) = e^{-r(\tau - t)}\left(f(t,T) - K \right),$$

at time t, where C and P denotes the price of a call and put option respectively.

Proof: The proof is similar to the arithmetic case. We get

$$\begin{aligned}
C(t,g) &= e^{-r(\tau-t)}\, \mathbb{E}[(f(\tau,T) - K)^+ | \mathcal{F}_t] \\
&= e^{-r(\tau-t)} \left(f(t,T) - K + \mathbb{E}[(K - f(\tau,T))^+ | \mathcal{F}_t] \right) \\
&= e^{-r(\tau-t)}(f(t,T) - K) + P(t,g),
\end{aligned}$$

from which the result follows. We have used that $t \mapsto f(t,T)$ is a martingale and the relation $(x-K)^+ = x - K + (K-x)^+$. ■

Hence, the price of a put option is:

Corollary 9.19 (†). *The time $t = 0$ price of a put option on an underlying asset that follows a geometric model with fixed delivery is given as*

$$e^{-rT}\left(K\Phi(-d) - f(0,T)\Phi(-(d+\xi))\right),$$

where d and ξ is as in Corollary 9.17.

Proof: Using the put call parity, we can easily rearrange and find that

$$\begin{aligned}
P(0) &= C(0) - e^{-rT}(f(0,T) - K) \\
&= e^{-rT}\left(f(0,T)\Phi(d+\xi) - K\Phi(d) - f(0,T) + K\right) \\
&= e^{-rT}\left(K(1-\Phi(d)) - f(0,T)(1-\Phi(d+\xi))\right) \\
&= e^{-rT}\left(K\Phi(-d) - f(0,T)\Phi(-(d+\xi))\right),
\end{aligned}$$

which was to be obtained. One could also do the same derivation as in Corollary 9.17. ■

We end this section with remark on the pricing of options of contracts with delivery periods. It is, to best of my knowledge, not possible to derive analytic expressions for options where the underlying commodity is delivered over a period, when modelling the driver as a geometric equation. This follows since the integral of the exponent of a general function $\int \exp(f)$ is not generally known. However, the payoff functions and the general setup will be much the same as in the arithmetic case. That is

$$p\left(\delta_{T_1-t}\mathcal{D}^{w}{}_{T_2-T_1}g(\tau)\right) = p\left(F(\tau, T_1, T_2)\right).$$

We can then take the discounted expectation of the representation above, but this has (at least to my knowledge) to be done numerically.

9.2 Lévy Models

We may also study the case where the forward dynamics are driven by a Lévy process instead of Wiener processes. Equation (75) then becomes

$$\mathrm{d}g(t) = \frac{\partial}{\partial x}g(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}L(t). \tag{93}$$

In (93), L is a zero mean and square integrable H-valued Lévy process, and σ is square integrable. In the same way as with the models driven by a Wiener process, we can find (using the results on stochastic differential equations in Hilbert spaces) show that

$$f(\tau) = S(\tau - t)g(t) + \int_t^{\tau} S(\tau - u)\sigma(u)\,\mathrm{d}L(u).$$

However, in order to compute the price of a forward contract $F(\cdot, T_1, T_2)$, we need to compute $\delta_{T_1-\tau}\mathcal{D}^w{}_{T_2-T_1} f(\tau)$. One can show in a similar fashion as earlier that this is

$$\delta_{T_1-\tau}\mathcal{D}^w{}_{T_2-T_1} f(\tau) = \delta_{T_1-t}\mathcal{D}^w{}_{T_2-T_1} g(t) + \int_t^\tau \delta_{T_1-u}\mathcal{D}^w{}_{T_2-T_1}\sigma(u)\,\mathrm{d}L(u).$$

However, in order to obtain more explicit representations, as with the earlier models, we need to introduce a class of functions known as *subordinated Wiener processes*. One can then show that

$$\int_t^\tau \delta_{T_1-u}\mathcal{D}^w{}_{T_2-T_1}\sigma(u)\,\mathrm{d}L(u) = \int_t^\tau \tilde{\sigma}(u)\,\mathrm{d}L(u),$$

where $\tilde{\sigma}$ is defined in exactly the same manner as previously. However, subordinated Wiener processes are beyond the scope of this thesis. We will therefore refer to Benth & Krühner [12, 11] for a more detailed discussion on such representations. When it comes to the issue of forward pricing, it is known that it is difficult to obtain closed form formulas when working with Lévy driven processes. One is therefore forced to use numerical methods, for example via Fourier techniques. As numerical methods are beyond the scope of this thesis as well, we refer to Benth, Benth & Koekebakker [10] and the references therein for a discussion on how to use Fourier techniques to price options in energy and related markets.

10 Estimating parameters and operators

In this section we will provide a discussion on the estimation of the parameters needed to model financial derivatives. We start by establishing ways to estimate the more simple parameters, like the mean, variance and correlation in a finite dimensional setting, before we establish some results on covariance operator estimation in Hilbert spaces.

10.1 Parameter estimation for finite dimensional derivatives

In finite dimensional spaces there are many methods of point estimation available to us. For example, the mean and variance may be estimated using maximum likelihood estimators. So if $\{X_i\}_{i=1}^n$ are i.i.d normal random variables, then

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

It is known that $\hat{\mu}$ is an unbiased estimator. That is, $\mathbb{E}[\hat{\mu}] = \mu$ (see e.g Devore & Berk [30]). As $\hat{\mu}$ is a sum of Gaussian random variables, $\hat{\mu}$ is itself Gaussian. We find easily that $\mathrm{Var}[\hat{\mu}] = \sigma^2/n$.[17]. Hence $\hat{\mu} \sim N(\mu, \frac{\sigma^2}{n})$.

The volatility σ, used in the various variations of the Black-Scholes formula, may also be estimated using maximum likelihood estimation, using an appropriate amount of historical data.[18] We find from any statistics textbook (e.g Devore & Berk [30]) that the maximum likelihood estimator for the variance is given as

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i - \hat{\mu})^2 .$$

It can be shown that $\mathbb{E}[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2$, hence $\hat{\sigma}^2$ is biased. It is however consistent.[19] In order for $\hat{\sigma}^2$ to be unbiased, we need to replace the factor $\frac{1}{n}$ with $\frac{1}{n-1}$, meaning that the unbiased then becomes

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \hat{\mu})^2 .$$

[17]On a sidenote, actually doing this computation has been shown to be quite difficult when the data is correlated. However, calculating the error of $\hat{\mu}$ is beyond the scope of this thesis, but is arguably best calculated by the automated blocking method, proposed by Flyvebjerg & Petersen [38], and proven by Jonsson [46]

[18]A much used alternative is by using the implied volatility instead of the historical volatility. See e.g Beckers [8] and Mayhew [54].

[19]An estimator in consistent if it converges in probability to the true value of the parameter.

Using the fact that if the sample is taking from a normal distribution, then $\frac{(n-1)\hat{\sigma}^2}{\sigma^2}$ is chi squared distributed with $n-1$ degrees of freedom, we get that

$$\mathrm{Var}\left[\frac{(n-1)\hat{\sigma}^2}{\sigma^2}\right] = \mathrm{Var}[\chi^2_{n-1}] = 2(n-1).$$

A simple rearrangement reveals $\mathrm{Var}[\hat{\sigma}^2] = \frac{2\sigma^4}{n-1}$. We know that all maximum likelihood estimators are asymptotically normal. However, assuming that the market data is normally distributed, is at best a gross oversimplification, so the distributional properties may not always hold. It was noted by Cizeau et al. [24] that the volatility is best described being log-normal.

Staying in the same framework, another parameter of interest is the correlation. This is especially import when dealing with spread options, see for example the earlier discussion of Margrabe's formula and its variations. The correlation ρ of two stochastic variables X and Y is defined by

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}[X]}\sqrt{\mathrm{Var}[Y]}}.$$

Since we are dealing with a sample of commodity prices, the sample correlation coefficient r_{xy}[20] is defined by

$$\begin{aligned} r_{xy} &:= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \\ &= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \\ &= \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}, \end{aligned} \tag{94}$$

which is called the Pearson's sample correlation coefficient.

Above, $\bar{x}, \bar{y}$ denotes the sample means of X and Y respectively, and s_x and s_y are the corrected sample standard deviations, i.e the same as (the corrected) $\hat{\sigma}$ above.

The expectation of r_{xy} must be approximated, and it can be shown that this expectation is given as $\mathbb{E}[r_{xy}] = \rho - \rho\frac{1-\rho^2}{2n} + ...$ For a more comprehensive treatment

[20] The sample correlation coefficient is denoted r_{xy} to avoid confusing it with the interest rate which is denoted r.

we refer to Fisher [36] and Zimmerman, Zumbo & Williams [76]. Following Bowley [16] and Hotelling [43] we find that $\mathrm{Var}[r_{xy}] = \frac{(1-r_{xy}^2)^2}{n}$.[21]

We can also find the density of r_{xy}. Following Olkin & Pratt [63], the density is given by

$$p(r_{xy}) = \frac{2^{n-2}}{\pi \Gamma(n-1)}(1-\rho^2)^{n/2}(1-r^2)^{(n-3)/2} \sum_{k=0}^{\infty} \Gamma^2\left(\frac{n+k}{2}\right)\frac{(2\rho r)^k}{k!}.$$

We refer to Olkin & Pratt for the derivation. They also propose an unbiased estimator for r_{xy} when the data follows a bivariate normal distribution. If we denote this estimator $\hat{\rho}$, they show that the asymptotic distribution of $\sqrt{n}(\hat{\rho}-\rho)$ is $N(0,(1-\rho^2)^2)$.

As a last trick in this short discussion on how to estimate the correlation coefficient, we introduce the Fisher transform, introduced by Fisher [36, 37]. The Fisher transformation states that if we define

$$z = \mathrm{atanh}(r_{xy}) = \frac{1}{2}\log(\frac{1+r_{xy}}{1-r_{xy}}),$$

then z is approximately normal with mean $\frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right)$ and variance $\frac{1}{n-3}$, for $n \neq 3$. Here n is the sample size and ρ is the true correlation coefficient. Following Bannör & Scherer [5] we find that the distribution of r_{xy} is approximately $N(\mathrm{atanh}(\rho_0), \frac{1}{n-3})$, where ρ_0 denotes the the true correlation coefficient.

10.2 Estimating the covariance operator

As we have seen in the previous subsection, the prices of most contingent claims in our framework depends on the covariance operator Q. However, in most cases, the covariance operator is not readily available to us, and must therefore be estimated. We will therefore focus on the covariance operator in this section. The covariance operator in this chapter is the covariance operator for a class of processes known as *autoregressive Hilbertian processes of order 1*. This section follows chapters 3 and 4 in Bosq [15] closely. Some results are from those chapters, while some are my own generalizations. We start by defining such processes.

[21] We note that there is some debate on this quantity. Several online resources state that $\mathrm{Var}[r_{xy}] = \frac{(1-\rho^2)}{n-2}$. I have however not managed to find any printed sources making this statement. See for example `http://www.sjsu.edu/faculty/gerstman/StatPrimer/correlation.pdf` or `http://strata.uga.edu/8370/lecturenotes/correlation.html`.

Definition 10.1 (Hilbert space valued white noise)**.** A sequence $\epsilon = \{\epsilon_n, n \in \mathbb{Z}\}$ of H-valued random variables is said to an H-valued *white noise* if

(i) $0 < \mathbb{E}[\|\epsilon_n\|^2] = \sigma^2 < \infty$, $\mathbb{E}[\epsilon_n] = 0$, $C_\epsilon := C_{\epsilon_n}$ do not depend on n

(ii) The sequence $\{\epsilon_n\}_{n\in\mathbb{Z}}$ is pairwise orthogonal.

If condition (ii) is replaced with

(iii) $\{\epsilon_n\}_{n\in\mathbb{Z}}$ is a sequence of i.i.d H-valued random variables,

then ϵ is said to be a *strong white noise*.

We can now define autoregressive Hilbertian processes of order 1, which from here on will be denotes $ARH(1)$-processes.

Definition 10.2 ($ARH(1)$-processes)**.** A sequence $X = \{X_n, n \in \mathbb{Z}\}$ of H-valued random variables is called an $ARH(1)$-process associated with (μ, ϵ, ρ) if it is stationary and for $n \in \mathbb{Z}$

$$X_n - \mu = \rho(X_{n-1} - \mu) + \epsilon_n,$$

where ϵ is an H-valued white noise, $\mu \in H$ and ρ is a continuous linear operator from H to H.

Remark 10.3. We note that the $ARH(1)$-processes are a generalization of the $AR(1)$-processes (autoregressive processes of order 1) used in (among others) Benth, Benth & Koekebakker [10] to model temperature derivatives.

From here on, let H be a separable Hilbert space, and $x \in H$. Also, we assume that the following hold

Condition 10.4. Let $X = \{X_n, n \in \mathbb{Z}\}$ be a standard $ARH(1)$-process associated with (ρ, ϵ) where ϵ is a strong white noise, and

$$\mathbb{E}\left[\|X_0\|^4\right] < \infty.$$

This condition will be referred to as *condition* A_1 from here on.

Then, we may define the covariance operator Q as in Bosq [15] by

$$Q(x) = \mathbb{E}[\langle X_0, x\rangle X_0]. \tag{95}$$

Now, if $\{X_i\}_{i=1}^n$ are observed, then a natural estimator of Q is

$$\hat{Q}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\langle X_i, x\rangle X_i, \;\; x \in H. \tag{96}$$

$\hat{Q}_n$ is called an *empirical covariance operator*. It is unbiased in the sense that $\mathbb{E}[\hat{Q}_n] = Q$, where the expectation is taken in $L_{HS}(H)$ which is the space of Hilbert-Schmidt operators on H.

Proposition 10.5. *The operator $\hat{Q}_n$ is a nuclear operator of finite range, meaning that $\hat{Q}_n$ is a Hilbert-Schmidt operator.*

Proof: See Bosq [15], chapter 4 and the references therein. ■

Moreover, if condition A_1 holds and $\|X_0\| < \infty$, then we have from **Theorem 4.2** in Bosq that

$$P\left(\|C - C_n\|_{L_{HS}(H)} > \epsilon\right) \leq 4\exp\left(-\frac{n\epsilon^2}{\alpha_1 + \beta_1\epsilon}\right),$$

and consequently, by choosing $\epsilon = A\left(\frac{\log n}{n}\right)^{1/2}$ and A such that $A^2 > \alpha_! + \beta_1 A$ we may apply the Borel-Cantelli lemma and conclude that

$$\|C_n - C\|_{L_{HS}(H)} = \mathcal{O}\left(\sqrt{\left(\frac{\log n}{n}\right)}\right)$$

with probability 1.This result and some more basic results on the upper bounds of $\|C_n - C\|_{L_{HS}(H)}$ given various conditions on $\|X_0\|$ can be found in chapter 4 in Bosq [15].

We will now turn our attention to estimation of the eigenelements of Q. If one is interested in data analysis of an observed process, then estimation of eigenvalues and eigenvectors is of great interest. These estimates may also be used to construct consistent estimators for ρ and other means of statistical predictions. We will from here on (in this subsection) use the notation from Bosq [15], however, all new notations will be explained. The results are also from Bosq, unless stated otherwise, and we will therefore refer to chapter 4.2 in the aforementioned monograph for proofs.

Being a compact self-adjoint operator, we may express Q using the following spectral decomposition [15][p. 102]

$$Q = \sum_{j=1}^{\infty} \lambda_j v_j \otimes v_j, \tag{97}$$

where $\{v_j\}$ is a complete orthonormal system in H and $\{\lambda_j\}$ is a sequence of real numbers such that $\sum_j \lambda_j < \infty$. We know from the spectral theorem that if these are eigenvalues, then $\lambda_j \to 0$. We have

$$Qv_j = \lambda_j v_j.$$

Bosq then introduces the natural estimators of these parameters, called the *empirical eigenelements* defined by

$$Q_n v_{j_n} = \lambda_{j_n} v_{j_n},$$

where $\{\lambda_{j_n}\}_j$ is a monotone decreasing sequence, and $\{v_{j_n}\}$ constitutes a complete orthonormal system in H.

If the eigenvectors $\{v_j\}$ are known, the natural estimators of the eigenvalues are defined as

$$\hat{\lambda}_{j_n} = \frac{1}{n} \sum_{i=1}^{n} \langle X_i, v_j \rangle^2, \tag{98}$$

where $j, n \geq 1$. These estimators are unbiased. We have

$$\mathbb{E}[\hat{\lambda}_{j_n}] = \mathbb{E}[\langle X_0, v_j \rangle^2] = \langle Q v_j, v_j \rangle = \lambda_j.$$

We note that the results regarding the asymptotics of the eigenvalues remain valid also when using the estimator in (98). This is proven in **Corollary 4.5** in Bosq [15], but essentially is because

$$\sup_{j \geq 1} |\hat{\lambda}_{j_n} - \lambda_j| \leq \|C_n - C\|_{L_{HS}(H)}.$$

The following results are from Bosq [15], and we need to introduce some notation. Bosq uses an alternative definition of the Signum function, and therefore we have that

$$\mathrm{sgn}(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ -1, & \text{otherwise} \end{cases} \tag{99}$$

He also defines the following vectors $v_{j_n} = \mathrm{sgn}\langle v_{j_n}, vj \rangle v_j$ for $j \geq 1$. He introduces this notation since v_j and $-v_j$ are both eigenvectors corresponding to the eigenvalue λ_j, it follows that v_j is *not* well defined as a statistical parameter, not even if the eigensubspace denoted $\mathcal{V}_j$ is one dimensional. Using this, we may state some consistency results for the empirical eigenvalues. The following result is **Theorem 4.4** in Bosq, which gives us some bounds on the eigenvalues of the covariance operator and its estimator. The proof in Bosq is very short, so I have written a new one.

Theorem 10.6. *Assume that condition A_1 holds true. Then*

$$\limsup_{n \to \infty} n \, \mathbb{E}\left[\sup_{j \geq 1} |\lambda_{j_n} - \lambda_j|^2 \right] \leq \sum_{h \in \mathbb{Z}} \mathbb{E}[\langle Z_0, Z_h \rangle_{L_{HS}(H)}], \tag{100}$$

and for all $\beta > \frac{1}{2}$ we have

$$\sqrt[4]{n}(\log n)^{-\beta} \sup_{j \geq 1} |\lambda_{j_n} - \lambda_j| \to 0 \tag{101}$$

with probability 1.

Additionally, if $\|X_0\| < \infty$, then

$$\sup_{j \geq 1} |\lambda_{j_n} - \lambda_j| = \mathcal{O}\left(\sqrt{\frac{\log n}{n}}\right) \tag{102}$$

with probability 1.

If one in addition assumes that $\mathbb{E}[\exp(\gamma\|X_0\|)^2] < \infty$ and equations (4.30) and (4.31) in Bosq [15], then it holds true that

$$\sup_{j \geq 1} |\lambda_{j_n} - \lambda_j| = \mathcal{O}\left(\frac{(\log n)^{5/2}}{\sqrt{n}}\right) \tag{103}$$

with probability 1.

Proof: We have from **Lemma 4.2** in Bosq that if ℓ_0 and ℓ_1 are compact linear operators with spectral decomposition

$$\ell_k = \sum_{j=1}^{\infty} a_{j,k} e_{j,k} \otimes f_{j,k}, \ k = 0, 1,$$

then $|a_{j,1} - a_{j,0}| \leq \|l_1 - l_0\|_{L(H)}$. Using this result compared with the spectral decomposition of Q given in (97) yields that

$$\sup_j |\lambda_{j_n} - \lambda_j| \leq \|C_n - C\|_{L(H)} \leq \|C_n - C\|_{L_{HS}(H)}. \tag{104}$$

It therefore follows that

$$\limsup_{n\to\infty} \mathbb{E}\left[\sup_j |\lambda_{j_n} - \lambda_j|^2\right] \leq \limsup_{n\to\infty} n\mathbb{E}\left[\|C_n - C\|^2_{L_{HS}(H)}\right] \to \sum_{h\in\mathbb{Z}} \mathbb{E}[\langle Z_0, Z_h\rangle_{L_{HS}(H)}],$$

by **Theorem 4.1**. By the same argument it also holds that

$$\sqrt[4]{n}(\log n)^{-\beta} \sup_j |\lambda_{j_n} - \lambda_j| \leq \sqrt[4]{n}(\log n)^{-\beta}\|C_n - C\|_{L_{HS}(H)} \to 0,$$

hence $\sqrt[4]{n}(\log n)^{-\beta} \sup_j |\lambda_{j_n} - \lambda_j| \to 0$ as $n \to \infty$. The last assertion follows in a similar way. ■

We may now focus on the estimation of the eigenvectors. In **Lemma 4.3**, Bosq proves that for some a_j defined as equations (4.45) and (4.46), we get that

$$\|v_{j_n} - v_j\| \leq a_j \|C_n - C\|_{L(H)}. \tag{105}$$

This bound allows us to use the asymptotic results concerning $\|C_n - C\|_{L(H)}$, so the following asymptotic result follows for the eigenvalues. The proof is again mine.

Theorem 10.7. *Assume that A_1 holds true. If $dim\mathcal{V}_j = 1$, then we have*

$$\limsup_{n\to\infty} \mathbb{E}[v_{j_n} - v_j]^2 \leq a_j^2 \sum_{h\in\mathbb{Z}} \mathbb{E}[\langle Z_0, Z_h\rangle]_{LHS(H)}, \tag{106}$$

and for all $\beta > \frac{1}{2}$,

$$\sqrt[4]{n}(\log n)^{-\beta} \|v_{j_n} - v_j\| \to 0 \tag{107}$$

with probability 1.

Additionally, if $\|X_0\| < \infty$, then

$$\|v_{j_n} - v_j\| = \mathcal{O}\left(\sqrt{\frac{\log n}{n}}\right) \tag{108}$$

with probability 1.

Finally, if $\mathbb{E}[\exp(\gamma\|X_0\|)^2] < \infty$ for some $\gamma > 0$ and X satisfies the same conditions as in the previous theorem, then

$$\|v_{j_n} - v_j\| = \mathcal{O}\left(\frac{(\log n)^{5/2}}{\sqrt{n}}\right) \tag{109}$$

with probability 1.

Proof: Applying equation (105), we find that

$$\limsup_{n\to\infty} \mathbb{E}[v_{j_n} - v_j]^2 \leq \limsup_{n\to\infty} n a_j^2 \, \mathbb{E}\left[\|C_n - C\|^2\right],$$

hence

$$\limsup_{n\to\infty} \mathbb{E}[v_{j_n} - v_j]^2 \leq a_j^2 \sum_{h\in\mathbb{Z}} \mathbb{E}[\langle Z_0, Z_h\rangle_{LHS(H)}].$$

Similarly,

$$\sqrt[4]{n}(\log n)^{-\beta} \leq \sqrt[4]{n}(\log n)^{-\beta}\|C_n - C\|_{L_{HS}(H)} \to 0,$$

hence

$$\sqrt[4]{n}(\log n)^{-\beta} \to 0.$$

The last assertion, follows (again) from a similar argument. ■

With respect to j is it possible to ensure uniform convergence the eigenvectors? This is possible for the eigenvalues, but clearly, eigenvectors are parameters who are much more sensitive to variation of operators than what the eigenvalues are. Bosq has the following result:

Corollary 10.8. *Assume that $\lambda_j \to 0$, and define for $k \geq 1$, $\Lambda_k := \sup_{1\leq j\leq k}(\lambda_j - \lambda_{j+1})^{-1}$. Then, if $\{k_n\}$ is a sequence of integers such that $\Lambda_{k_n} = o(\sqrt{n})$ we get*

$$\mathbb{E}\left[\sup_{1\leq j\leq k_n} \|v_{j_n} - v'_{j_n}\|^2\right] \to 0. \tag{110}$$

Note that if there exists a convex function φ such that $\varphi(j) = \lambda_j$ for $j \geq 1$, then $\Lambda_k = (\lambda_k - \lambda_{k+1})^{-1}$.

Moreover, if $\|X_0\| < \infty$ and $\Lambda_{k_n} = o\left(\sqrt{\frac{n}{\log n}}\right)$, then

$$\sup_{1\leq j\leq k_n} \|v_{j_n} - v'_{j_n}\| \to 0 \tag{111}$$

with probability 1.

Proof: This is **Corollary 4.3** in Bosq's monograph. ■

Given that $\mathcal{V}_j$ is not one dimensional, we may construct a result on the upper bound for the eigenvector in the same manner as with the eigenvalues. If we let Π_a^b denote the orthogonal projector of the eigensubspace $\mathcal{V}_a^b$, which is generated by $\{v_a, ..., v_b\}$, for $v_a, ..., v_b$ associated with the eigenvalue $\lambda_a (= \lambda_b)$, then for some c_j defined in **Lemma 4.4** and $j \in \{a, ..., b\}$ we have that

$$\|v_{j_n} - \Pi_a^b(v_{j_n})\| \leq c_j\|C_n - C\|_{L(H)}. \tag{112}$$

From this it is possible to construct results like the two previous results but where $\dim\mathcal{V}_j > 1$.

Theorem 10.9 (†). *Assume condition A_1 holds. Then*

$$\limsup_{n\to\infty} \mathbb{E}\left[\|v_{j_n} - \Pi_a^b(v_{j_n})\|^2\right] \le c_j^2 \sum_{h\to\mathbb{Z}} \mathbb{E}[\langle Z_0, Z_h\rangle_{LHS(H)}],$$

and for all $\beta > \frac{1}{2}$, it holds true that

$$\sqrt[4]{n}(\log n)^{-\beta}\|v_{j_n} - \Pi_a^b(v_{j_n})\| \to 0$$

with probability 1.

Finally, if X satisfies $\alpha_k \le ar^k$ for $0 \le r \le 1$, $\alpha > 0$ and $k \ge 1$, and if X satisfies $\lambda_j^Z \le ar^j$ for $j \ge 1$ and $\mathbb{E}[\exp(\gamma\|X_0\|^2)] < \infty$ for some $\gamma > 0$, then

$$\|v_{j_n} - \Pi_a^b(v_{j_n})\| = \mathcal{O}\left(\frac{(\log n)^{5/2}}{\sqrt{n}}\right)$$

with probability 1.

Proof: Using equation (112) and (104) we find that $\|v_{j_n} - \Pi_a^b(v_{j_n})\| \le \|C_n - C\|_{LHS(H)}$. Hence

$$\limsup_{n\to\infty} \|v_{j_n} - \Pi_a^b(v_{j_n})\|^2 \le c_j^2 \limsup_{n\to\infty} n\|C_n - C\|_{LHS(H)}^2,$$

and therefore

$$\limsup_{n\to\infty} \mathbb{E}\left[\|v_{j_n} - \Pi_a^b(v_{j_n})\|^2\right] \to \sum_{h\in\mathbb{Z}} \mathbb{E}[\langle Z_0, Z_h\rangle_{LHS(H)}].$$

In the same manner we find that

$$\sqrt[4]{n}(\log n)^{-\beta}\|v_{j_n} - \Pi_a^b(v_{j_n})\| \le c_j^2 \sqrt[4]{n}(\log n)^{-\beta}\|C_n - C\|_{LHS(H)},$$

which proves that

$$\sqrt[4]{n}(\log n)^{-\beta}\|v_{j_n} - \Pi_a^b(v_{j_n})\| \to 0.$$

The last assertion follows from (104), and the results concerning the bound on $\|C_n - C\|_{LHS(H)}$. ∎

We conclude this section with some results on the distributions of the estimators described.

Theorem 10.10. *In $L_{HS}(H)$ we have that*

$$\sqrt{n}(C_n - C) \overset{\mathcal{D}}{\to} N_1 \sim N(0, \Gamma_1), \tag{113}$$

where Γ_1 satisfies equation (4.88) in Bosq.

Proof: This is **Corollary 4.6** in Bosq. ■

Theorem 10.11. *If A_1 holds true and $\mathcal{V}_j$ is one dimensional, then*

$$\sqrt{n}(\lambda_{j_n} - \lambda_j) \overset{\mathcal{D}}{\to} N_2 \sim N(0, \sigma_j^2), \tag{114}$$

where σ_j^2 is given by (4.90) in Bosq. This result also holds true if we replace λ_{j_n} with $\hat{\lambda}_{j_n}$.

Moreover,

$$\sqrt{n}(v_{j_n} - v'_{j_n}) \overset{\mathcal{D}}{\to} N_3 \sim N(0, \Gamma_2), \tag{115}$$

where

$$\Gamma_2 = \left(\sum_{k \neq j} (\lambda_j - \lambda_k)^{-1} v_k \otimes v_k \right) [N_1(v_j)].$$

Proof: This is **Theorem 4.10**, **Corollary 4.7** and **Corollary 4.8** in Bosq. ■

11 Model uncertainty

As mentioned in the introduction, there are several examples where market participants has taken huge losses due to model uncertainty. In this chapter we will define what model uncertainty is, some of the sources of model uncertainty and some methods to measure and quantify this uncertainty.

11.1 Background on model uncertainty

We start by defining what model uncertainty is. This definition is from the framework of Cont [25] and Bannör & Scherer [5].

Definition 11.1 (Model uncertainty)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{Q}$ be a family of equivalent martingale measures on $(\Omega, \mathcal{F})$, such that every discounted asses price is a Q-martingale for all $Q \in \mathcal{Q}$. If $|\mathcal{Q}| > 1$ we have model uncertainty, where $|\mathcal{Q}|$ denotes the cardinality of $\mathcal{Q}$. Furthermore, denote by $\mathcal{P}$ the set of all objective probability measures on Ω.

Definition 11.2 (Parameter uncertainty)**.** Let $\{P_\theta\}_{\theta \in \Theta}$ be a family of pairwise different probability measures on $(\Omega, \mathcal{F})$ such that the models $(\Omega, \mathcal{F}, \mathcal{F}_t, P_\theta)$ are arbitrage free for all $\theta \in \Theta$. The model faces *parameter uncertainty* if $|\Theta| > 1$, where $|\Theta|$ denotes the cardinality of Θ.

What are the sources of model uncertainty? One of the earliest papers on model uncertainty was published by Derman [29], and in this paper he identifies several sources contributing to model uncertainty in a financial context. These sources were further generalized by Gupta, Reisinger & Whitley [67], and we summarize them here:

1. Incorrect model: Mathematical models may fail to project the movements of a stock (or some other commodity). Some factors may have been forgotten or incorrectly modelled - for example modelling the volatility of the interest rate as deterministic (or even constant) when they should have been stochastic, or vice versa. Moreover, even though a model may perform well in a calm market, in may become inappropriate when the markets face severe instability, for example in connection with a financial crisis. Also, some models assume no transaction costs (Black-Scholes-Merton for example), and these models may break down in a market with transaction costs and low liquidity.

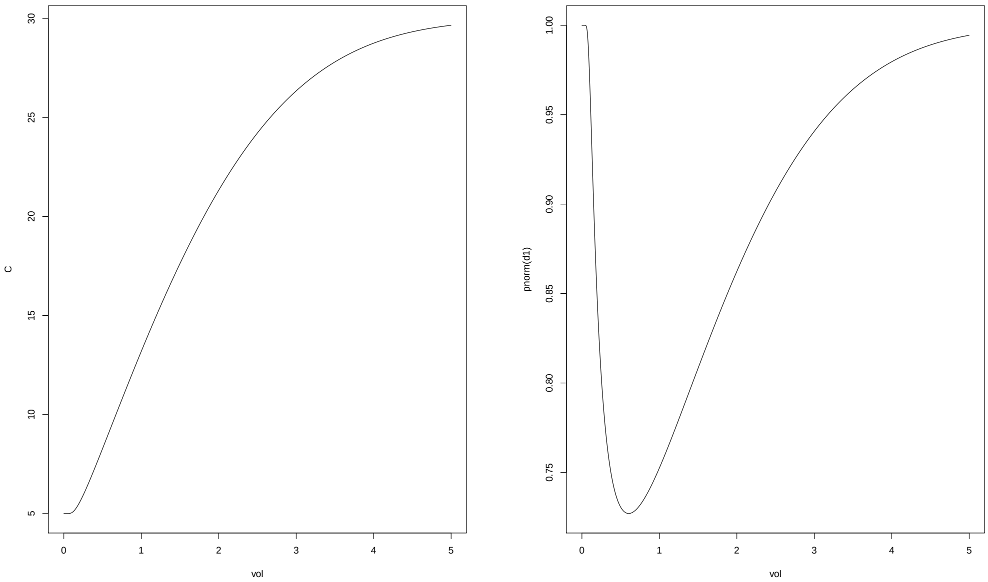
2. Incorrect solution: Having the correct model does not help if the solution is wrong. For example, Li's copula formula [50] has a missing closing

bracket[22]. Other similar sources may be faulty programming, either a actual error in the software, as for example the R2009b release of Matlab, or that the market participant forgot to declare a float point number the correct way.

3. Incorrect calibration: If the underlying process is not stationary, it may cause the earlier parameter calibrations to become erroneous. The same may happen if the solutions are unstable - which may lead to the choice of wrong model. Furthermore, there may be a lack of robustness of the solutions with respect to the modelling assumptions.

Example 11.3. If there is uncertainty on the value of the volatility of a stock, we may get a dramatic increase of the price of a call option in the Black-Scholes framework, since the option price is monotone with respect to its volatility. Using that the spot price is $S(t) = 30$, with strike price $K = 25$ and maturity time $T = 1$, we find that the price may become six times greater than for an asset whose volatility is zero!. It also has a dramatic effect on the delta hedge. We notice that when $\sigma \to \infty$ the price of the call option tends to $S(t)$ and when $\sigma \to 0$ the price tends to $S(t) - K$. These properties will be proven later.
See figure 1 for a graphical explanation.

[22]The source of this statement is Gupta et al [67]. They also claim that billions of dollars have been invested following this formula and its error.



(a) Price (b) Delta hedge

Figure 1: This figure shows how an increase in volatility dramatically alters the price (figure (a)) and delta hedge (figure (b)) of a call option in the Black-Scholes framework. The figures shows the price and the delta hedge as functions of the underlying assets volatility.

Based on the example above, we see that we may prove the following model-free elementary bounds on the price of call and put options.

Proposition 11.4. *Let $C(t)$ and $P(t)$ denote the price of a call option and put option respectively. Then $C(t) \leq S(t)$ and $P(t) \leq Ke^{-rt}$, where the notation is the same as in the Black-Scholes formula. Both prices are bounded below by 0.*

Proof: Knowing the Black-Scholes formula is monotonic as function of its volatility (see appendix for proof), we must have that it reaches its maximum when the volatility tends to infinity. We know from Black & Scholes [14] that

$$C(t) = S(t)\Phi(d_1) - Ke^{-rt}\Phi(d_2)$$

and

$$P(t) = Ke^{-rt}\Phi(-d_2) - S(t)\Phi(-d_1),$$

where all parameters are as they always are.

We find that

$$\lim_{\sigma\to\infty} d_1 = \infty, \quad \lim_{\sigma\to\infty} d_2 = -\infty,$$

which implies that $\Phi(d_1)$ tends to 1 and $\Phi(d_2)$ tends to 0. This shows that $C(t) \leq S(t)$. Reversing the signs proves the statement for the put options.
The payoff function for calls and puts are defined as $p_c(x) = \max(0, x-K)$ and $p_p(x) = \max(0, K-x)$ respectively, from which the lower bound follows. ■

Remark 11.5. We note that it can be found in Chance [21] that if $C(t)^A$ and $P(t)^A$ denotes the prices of American options, then the upper bounds are $C(t)^A \leq S(t)$ and $P(t)^A \leq K$. From this it is easy to prove that $C(t) \leq C(t)^A$ and $P(t) \leq P(t)^A$ since $e^{-rt} \leq 1$. Chance also shows that the price of an American call option is always higher than its *intrinsic value*. That is, $C(t)^A \geq \max(0, S(t)-K)$. This may be proved by simple no-arbitrage arguments. They also show that the lower bound for a European call option and therefore also for an American call option is $\max(0, S(t) - Ke^{-rt})$ rather than just zero. A reversal of the signs yields the lower bounds for put options. We refer to Chance for a further discussion on such elementary model-free bounds on vanilla options.[23]

11.2 Risk measures

In this subsection we will define the notion of *risk measures*. Measuring risk is vital to the financial industry. As we will see, depending on the uncertainty on certain parameters, the price of illiquid derivatives may vary greatly, and thus leading to mispriced derivatives, which in some extreme cases may lead to bankruptcies and collapses. Risk measures are used to quantify the risk faced by an asset or more generally a portfolio. Since they are expressed in monetary units, we will also argue that they may be used to calculate the reserve needed to make a financial position acceptable. We will define both *coherent* and *convex* risk measures, in the frameworks of several authors. Some commonly used measures of risk are called *Average value-at-risk* and *Expected shortfall*[24], which quantifies potential extreme losses in the tail of of the distribution of the possible returns.
We give the definition of a coherent risk measure as defined by Artzner et al. [2].

[23] The findings in Chance's paper, is completely consistent with the results in part three of Margrabe's [53] seminal 1978 paper, where he finds that a European type exchange option (spread option) is *"...worth more alive than dead"* - implying that a trader should wait to the very last minute to exercise the option, proving that the formula is also valid for American options. Moreover, the minimum value $S(t) - Ke^{-rt}$ is greater than the value if exercised, which is $S(t) - K$. Which then implies that the option is worth *more* by simply selling it in the market!

[24] Also known as *Conditional value-at-risk.*

Definition 11.6 (Coherent risk measure)**.** Define the payoff of a claim as a bounded and measurable function $X : \Omega \to \mathbb{R}$ defined on the set Ω of market scenarios and let E denote the set of payoffs. Then, a *coherent risk measure* is a map $\rho : E \to \mathbb{R}$ such that:

1. Monotonicity: If a portfolio X dominates another portfolio Y in terms of payoffs then it should be less risky. I.e, $X \geq Y \implies \rho(X) \leq \rho(Y)$.

2. Risk is measured in monetary units. Adding to a portfolio X a sum a in numéraire reduces risk by a. That is, $\rho(X + a) = \rho(X) - a$.

3. Subadditivity, diversification reduces risk: $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

4. Positive homogeneity, the risk of a position is proportional to its size: $\rho(\lambda X) = \lambda \rho(X)$, for all $\lambda \geq 0$.

The notion of coherent risk measures has been generalized by Föllmer & Schied [40] into *convex risk measures*.

Definition 11.7 (Convex risk measure)**.** Replace condition 3 and 4 with a convexity condition. That is,

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y),$$

for all $0 \leq \lambda \leq 1$.

Remark 11.8. To gain some intuition, it is common to interpret ρ as follows: It is stated in Cont [25] that a risk measure is defined in monetary units. In other words, $\rho(X)$ is the amount of money that should be added to a portfolio in a risk free way to make a financial position acceptable for a given risk tolerance for a trader. The convexity condition is roughly speaking just a reflection of the notion that diversification reduces risk. The second condition in the definition, $\rho(X + a) = \rho(X) - a$, can be interpreted as follows. If one adds an amount a to a portfolio in a risk free way, the capital requirement to make the portfolio's risk lie in the risk tolerance of the trader should be reduced by the same amount. It is stated in the introduction of Artzner et al. [2] that "*...these measures of risk can be used as (extra) capital requirements to regulate the risk assumed by market participants, traders, and insurance underwriters, as well as to allocate existing capital*".

Remark 11.9. It should be noted that there exists several different definitions of a risk measure. The monotonicity definition used in this dissertation is actually *anti-monotone*, meaning that $X \leq Y \implies \rho(Y) \leq \rho(X)$, in order to be consistent with notion of using risk measures to calculate the reserve needed to make a financial position acceptable. There are some other authors, for example Rockafellar [68], that uses monotonicity in its more known form, i.e $X \leq Y \implies \rho(X) \leq \rho(Y)$.

From remark 11.8, it is clear that since $\rho(X)$ represents the amount of money needed to make a position acceptable, one must wish that $\rho(X) \leq 0$. This leads to the notion of the *acceptance set* of ρ. The acceptance set is the set of all positions X such that this condition is satisfied. In mathematical notation this is:

Definition 11.10. Let $(\Omega, \mathcal{F})$ be a measurable space. Let $\mathbb{V}$ denote the space of all measures on $(\Omega, \mathcal{F})$ and let $\mathbb{X}$ denote a linear space of functions $X : \Omega \to \mathbb{R}$, in which the constant functions are contained.

Definition 11.11 (Acceptance set of a risk measure)**.** Any risk measure $\rho : \mathbb{X} \to \mathbb{R}$ induces an *acceptance set* denoted $\mathcal{A}_\rho$, defined as

$$\mathcal{A}_\rho = \{X \in \mathbb{X} : \rho(X) \leq 0\}. \tag{116}$$

The above definition shows that given a risk measure, we can define an acceptance set. The next definition shows that given an acceptance set, we can define a risk measure on it.

Definition 11.12. Let $\mathcal{A} \in \mathbb{X}$ be a set of acceptable random variables, that is, $X \in \mathcal{A} \implies \rho(X) \leq 0$. Then the set $\mathcal{A}$ has an associated risk measure $\rho_{\mathcal{A}}$ defined as follows:

$$\rho_{\mathcal{A}}(X) = \inf\{m \in \mathbb{R} : m + X \in \mathcal{A}\}. \tag{117}$$

Then, it is possible to prove the following set of useful properties of the aforementioned risk measures. We note that the following theorem is a generalization of **Proposition 2** in Föllmer & Schied [40]. The generalization is done in **Theorem 2.17** in Dahl [26].

Theorem 11.13. *Let ρ be a convex risk measure with induced acceptance set $\mathcal{A}_\rho$. Then*

(i) $\rho_{\mathcal{A}_\rho} = \rho$

(ii) $\mathcal{A}_\rho$ *is a nonempty convex set*

(iii) If $X \in \mathcal{A}_\rho$ and there exists Y such that Y dominates X, then $Y \in \mathcal{A}_\rho$

(iv) ρ is a coherent risk measure only if $\mathcal{A}_\rho$ is a convex cone

Conversely, let $\mathcal{A}$ be a nonempty convex subset of $\mathbb{X}$. Let $\mathcal{A}$ be such that if $X \in \mathcal{A}$ and $Y \in \mathbb{X}$ such that Y dominates X implying $Y \in \mathcal{A}$, then the following holds true

(v) $\rho_{\mathcal{A}}$ *is a convex risk measure*

(vi) $\mathcal{A}$ *is a convex cone only if* $\rho_{\mathcal{A}}$ *is a coherent risk measure*

(vii) $\mathcal{A} \subset \mathcal{A}_{\rho_{\mathcal{A}}}$.

Proof: As (i), (ii) and (v)-(vii) is proven by Dahl [26], but she left the proofs of properties (iii) and (iv) to the reader. We will therefore only prove these two properties.

(iii) We have to prove that $\rho(Y) \leq 0$. We know that from the definition of a risk measure that $X \leq Y$ only if $\rho(X) \geq \rho(Y)$. Hence, $\rho(Y) \leq \rho(X) \leq 0$, which implies that $Y \in \mathcal{A}_\rho$.

(iv) Assume $X, Y \in \mathcal{A}_\rho$ and that $\alpha, \beta \geq 0$. Then, since ρ coherent, we have by positive homogeneity and subadditivity that $\rho(\alpha X) = \alpha\rho(X)$ and $\rho(X + Y) \leq \rho(X) + \rho(Y)$. And therefore, if $\rho(X) \leq 0$ and $\rho(Y) \leq 0$, then

$$\rho(\alpha X + \beta Y) \leq \alpha\rho(X) + \beta\rho(Y) \leq 0,$$

proving that if $X, Y \in \mathcal{A}_\rho$, then $\alpha X + \beta Y \in \mathcal{A}_\rho$ for any $\alpha, \beta \geq 0$, which proves that $\mathcal{A}_\rho$ is a convex cone.

■

Convex risk measures have several representations. We will state two of them.

Theorem 11.14. *Let $\rho : \mathbb{X} \to \mathbb{R}$ be a convex risk measure and $V \in \mathbb{V}$. Assume in addition that ρ is lower semicontinuous. Then $\rho = \rho^{**}$. Hence for each $X \in \mathbb{X}$*

$$\begin{aligned}\rho(X) &= \sup\{\langle X, v\rangle - \rho^*(v) : v \in V\} \\ &= \sup\{\langle X, v\rangle - \rho^*(v) : v \in \mathrm{dom}(\rho^*)\},\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ is a pairing between $\mathbb{X}$ and V.

Proof: See Dahl **Theorem 2.2.1** in [26] and the references therein. ■

With this knowledge, we have from **Theorem 5** Föllmer & Schied [40] that we have the following representation of ρ by restricting Ω to be finite dimensional.

Theorem 11.15. *Assume that $\mathbb{X}$ is the space of all real-valued functions on a finite state space Ω, and denote by $\mathcal{P}$ the set of all probability measures on Ω. Then $\rho : \mathbb{X} \to \mathbb{R}$ is a convex risk measure if and only if there exists a penalty function $\alpha : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$ such that*

$$\rho(X) = \sup_{Q \in \mathcal{P}} \left\{ \mathbb{E}_Q[-X] - \alpha(Q) \right\}, \tag{118}$$

where $\mathbb{E}_Q[\cdot]$ denotes the expectation with respect to $Q \in \mathcal{P}$. The function α satisfies $\alpha(Q) \geq -\rho(0)$ for any $Q \in \mathcal{P}$, and is both convex and lower semicontinuous on $\mathcal{P}$.

Proof: See Föllmer & Schied [40], **Theorem 5**. ■

Remark 11.16. It is stated in Cont [25] and proven in Lüthi & Doege [31] that if ρ is a coherent risk measure, then α in (118) only takes the values 0 and infinity. Zero if $Q \in \mathcal{P}$ and infinity otherwise.

Remark 11.17. We can find in the proof of Föllmer & Schied's **Theorem 5** that α can be expressed as

$$\alpha(Q) = \sup_{X \in \mathcal{A}_\rho} \left\{ \mathbb{E}_Q[-X] \right\}.$$

Dahl [26] has a different proof of the same result, and shows that $\alpha(Q) = \rho^*(-Q)$ for all $Q \in \mathcal{P}$, where ρ^* denotes the convex conjugate of ρ. The same result is also found in Lüthi & Doege [31].

In the subsequent chapter, we will show that (118) holds in infinite dimensions as well. The representations above, provide us with good understanding of risk measures on the real line (extendable to the n-dimensional plane), in the sense that any convex risk measure has the representation

$$\rho(X) = \sup_{Q \in \mathcal{P}} \left\{ \mathbb{E}_Q[-X] - \sup_{X \in \mathcal{A}_\rho} \mathbb{E}[-X] \right\}.$$

In [25], Cont defines an axiomatic setting for model uncertainty. Consider a measurable space $(\Omega, \mathcal{F})$ of market scenarios, without a reference probability. The underlying assets is denoted as a measurable mapping $S : \Omega \to D([0, T])$, where $D([0, T])$ denotes the space of càdlàg functions, and $S(\omega)$ denotes the trajectory of the price in the market scenario $\omega \in \Omega$. Any claim on S will be represented by a random variable H. All assets and payoffs are assumed to be discounted values.

Definition 11.18 (Benchmark instruments)**.** Options written on an instrument S whose prices are observed on the market are called *benchmark instruments* or *benchmark options*. Their prices are denoted by $(C_i^*)_{i\in I}$ and the payoffs by $(H_i)_{i\in I}$, where I denotes some index set. Since the markets are not complete, a unique price is most often not available - instead we have a range[25] of prices. Hence $C_i^* \in [C_i^{\text{bid}}, C_i^{\text{ask}}]$ to accommodate for the bid-ask spread.

Moreover, it is also assumed that there exists a set of arbitrage-free pricing models $\mathcal{Q}$ replicating the market observed prices of the benchmark, $S(t)$ is a martingale under Q for each $Q \in \mathcal{Q}$ and $0 \le t \le T$, $\mathbb{E}_Q[|H_i|] < \infty$ and $\mathbb{E}_Q[H_i] = C_i^*$. The last condition may be relaxed in the way that $\mathbb{E}_Q[H_i] \in [C_i^{\text{bid}}, C_i^{\text{ask}}]$.

We can now state Cont's axioms of what a measure on model uncertainty should verify[26]

1. There is no model uncertainty on the value of a liquid option.
2. If an option can be (partially) hedged in a model free way, this should reduce the model uncertainty. If the option can be totally hedged in a model free way, then there is no model uncertainty on its value.
3. Liquid options may be used to hedge more complex instruments.
4. The model uncertainty on the value of a portfolio should be expressed in monetary units and be normalized to make it comparable to the value of the portfolio.
5. Diversification should lead to a decrease in the model uncertainty on the value of a portfolio.

These requirements are then formalized mathematically. Let X be a contingent claim, and consider a mapping $\mu : \mathcal{C} \mapsto [0, \infty]$, where $\mathcal{C}$ denotes the set of contingent claims with well-defined prices in all models. That is,

$$\mathcal{C} = \left\{ H \in \mathcal{F}_T, \sup_{Q\in\mathcal{Q}} \mathbb{E}_Q[|H|] < \infty \right\}.$$

We can now state the requirements above:

[25]known as the bid-ask spread, which is the amount of money by which the ask price exceeds the bid price.

[26]This is only a short summary, we refer to Cont [25] chapter 3.3 for a more comprehensive list.

1. For all liquid benchmark instruments, model uncertainty reduces to the uncertainty on the market value. Hence

$$\mu(H_i) \leq |C_i^{\text{ask}} - C_i^{\text{bid}}|,$$

where $i \in I$.

2. If we hedge on the underlying, then we get the following effect

$$\mu\left(X + \int_0^T \phi(t)\,\mathrm{d}S(t)\right) = \mu(X).$$

In particular, the value of an instrument that can be hedged in a model free way, has no model uncertainty, meaning that if $Q(X = x + \int_0^T \phi(t)\mathrm{d}S(t)) = 1$, then $\mu(X) = 0$. Here, for a simple predictable process $\{\phi(t), 0 \leq t \leq T\}$ representing a self-financing hedging strategy $\int_0^x \phi(t)\,\mathrm{d}t$ (whose discounted value is a martingale under the measure Q) represents the gain from trading from 0 to x.[27]

3. From the convexity of the map μ, it follows that model uncertainty decreases though diversification. Let $t \in [0, 1]$, then

$$\mu(tX_1 + (1-t)X_2) \leq t\mu(X_1) + (1-t)\mu(X_2).$$

4. Hedging with liquidly traded options:

$$\mu\left(X + \sum_{i=1}^{k} u_i H_i\right) \leq \mu(X) = \sum_{i=1}^{k} |u_i(C_i^{\text{bid}} - C_i^{\text{ask}})|,$$

and in particular, if X can be statically replicated by liquid options, then the model uncertainty reduces to the uncertainty of the replications costs. Meaning

$$\mu(X) \leq \sum_{i=1}^{k} |u_i||C_i^{\text{bid}}, C_i^{\text{ask}}|.$$

[27] The integral $\int_0^T \phi(t)\,\mathrm{d}S(t)$ is defined to be the following: We can construct a stochastic integral with respect to $\mathcal{Q}$ such that for each $Q \in \mathcal{Q}$ there exists a process $G_t(\phi)$ such that $G_t(\phi) = \int_0^T \phi(t)\,\mathrm{d}S(t)$ Q-almost surely. According to Cont, this is done in the framework of the article **Intégrales stochastiques par rapport à une famille de probabilités** by Doléans-Dade. We refer the interested reader to that article for more information.

With these ingredients, Cont then constructs two measures of model uncertainty. First a coherent measure, then a convex measure. These are both based on the "worst case approach", which is an approach where assumes the largest difference among the set of prices, hence the name.

For a payoff $X \in \mathcal{C}$, the upper and lower price bounds are defined by

$$\bar{\pi}(X) := \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q[X] \text{ and } \underline{\pi}(X) := \inf_{Q \in \mathcal{Q}} \mathbb{E}_Q[X]. \tag{119}$$

It is noted by Gupta et al [67] that the lower bound $\underline{\pi}(X)$ is the conservative bid price.

The mapping $X \mapsto \bar{\pi}(-X)$ then defines a coherent risk measure, and any of the pricing models $Q \in \mathcal{Q}$ will ensure that the value of X falls in the interval $[\underline{\pi}(X], \bar{\pi}(X)$. Clearly, if there is no model uncertainty on the value of X, then $\bar{\pi}(X) = \underline{\pi}(X)$. In **Proposition 4.1**, Cont [25] proves that for any benchmark derivative, we have

$$C_i^{\text{bid}} \leq \underline{\pi}(H_i) \leq \bar{\pi}(H_i) \leq C_i^{\text{ask}} \tag{120}$$

and $\mu_{\mathcal{Q}} : \mathcal{C} \to \mathbb{R}^+$ defined by

$$\mu_{\mathcal{Q}}(X) := \bar{\pi}(X) - \underline{\pi}(X) \tag{121}$$

constitutes a measure on model uncertainty verifying the properties outlined above.

We will now present some clarifying examples.

Example 11.19. Using the bounds in Proposition 11.4 and the subsequent remark, we find that for a European call option, $\bar{\pi}(X) = S(t)$ and $\underline{\pi}(X) = \max(0, S(t) - Ke^{-rt})$. Hence $S(t) - \max(0, S(t) - Ke^{-rt})$ is a measure for the model uncertainty on a portfolio consisting of a single call option. □

Another intuitive example to gain some intuition, can be constructed as follows:

Example 11.20. Consider the price approximation on spread options where the underlying follows a geometric Brownian motion, as done by Carmona & Durrleman [17] or in section 5.6 in this thesis. In another article [18], Carmona & Durrleman derive upper and lower bounds for this approximation. For example, they prove in **Proposition 11** that we have the following upper bound

$$\begin{aligned} \Pi \leq \min \Bigg\{ & \min(\bar{\Pi}(0), \overline{\Pi \circ s}(0)) + \alpha \Psi(\beta \rho'), \min(\bar{\Pi}(-\phi), \overline{\Pi \circ t}(0)) + \gamma \Psi(\delta \rho'), \\ & \min(\overline{\Pi \circ s}(-\psi), \overline{\Pi \circ t}(\psi - \pi)) + |\kappa| \Psi\left(\frac{\beta \delta \rho'}{\sigma}\right) \Bigg\}, \end{aligned} \tag{122}$$

and a lower bound $\hat{\Pi} \leq \Pi$ given as

$$\hat{\Pi} = \sup_{\theta \in \mathbb{R}} \left\{ \sup_{d \in \mathbb{R}} [a\Phi(d + \beta \cos(\theta + \phi)) - \gamma \Phi(d + \delta \cos(\theta)) - \kappa \Phi(d)] \right\}. \tag{123}$$

If we denote the upper bound by $\bar{\pi}$ and the lower bound by $\underline{\pi}$, then by construction the price Π (or in many cases the approximation of the price) will lie in the interval $[\underline{\pi}, \bar{\pi}]$. We refer to Carmona & Durrleman [18] for the notation. □

To illustrate even further, we give two more examples. The first one being on the Black-Scholes formula for put options, and second on Margrabe's formula.

Example 11.21. Consider a put option on a commodity $S(t)$, whose dynamics is given as a geometric Brownian motion with non-constant, time-dependent volatility:

$$\mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \sigma_i(t)S(t)\,\mathrm{d}W(t). \tag{124}$$

In (124), $\{W(t), t \geq 0\}$ is a Brownian motion and σ_i is a model dependent volatility, meaning that for each $Q_i \in \mathcal{Q}$ there exists a corresponding volatility $\sigma_i(\cdot)$.

We know from [14] that the price of a call option option is

$$V(t, T, K, S(t), \Sigma) = S(t)\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2), \tag{125}$$

where Σ is the implied volatility and

$$d_1 = \frac{\log(S(t)/K) + \left(r + \Sigma^2/2\right)(T-t)}{\Sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \Sigma\,(T-t)\,.$$

From the put-call parity we can easily deduce that the fair arbitrage free price of a put option is

$$P(t, T, K, S(t), \Sigma) = -S(t)\Phi(-d_1) + Ke^{-rt}\Phi(-d_2), \tag{126}$$

where all the quantities are the same as in the call option case.

A simple application of Itô's formula reveals that the solution of (124) is

$$S(t) = S(0)\exp\left(rt - \frac{1}{2}\int_0^t \sigma_i^2(s)\,\mathrm{d}s + \int_0^t \sigma_i(s)\,\mathrm{d}W(s)\right). \tag{127}$$

Define $\hat{\sigma}^2 = \frac{1}{t}\int_0^t \sigma_i^2(s)\,\mathrm{d}s$, and insert this into (127), and we get:

$$S(t) = S(0)e^{(r - \frac{1}{2}\hat{\sigma}_i^2(t))t + \int_0^t \sigma_i(s)\,\mathrm{d}W(s)}. \tag{128}$$

We know that the solution of a geometric Brownian motion that generates the Black & Scholes formula is

$$S(t) = S(0)e^{(r-\frac{1}{2}\Sigma^2)t+\Sigma W(t)}, \tag{129}$$

so by comparing (128) and (129) we see that the calibration condition is simply

$$\frac{1}{t}\int_0^t \sigma_i^2(s)\,\mathrm{d}s = \Sigma^2. \tag{130}$$

(130) has many solutions, one proposed by Cont is

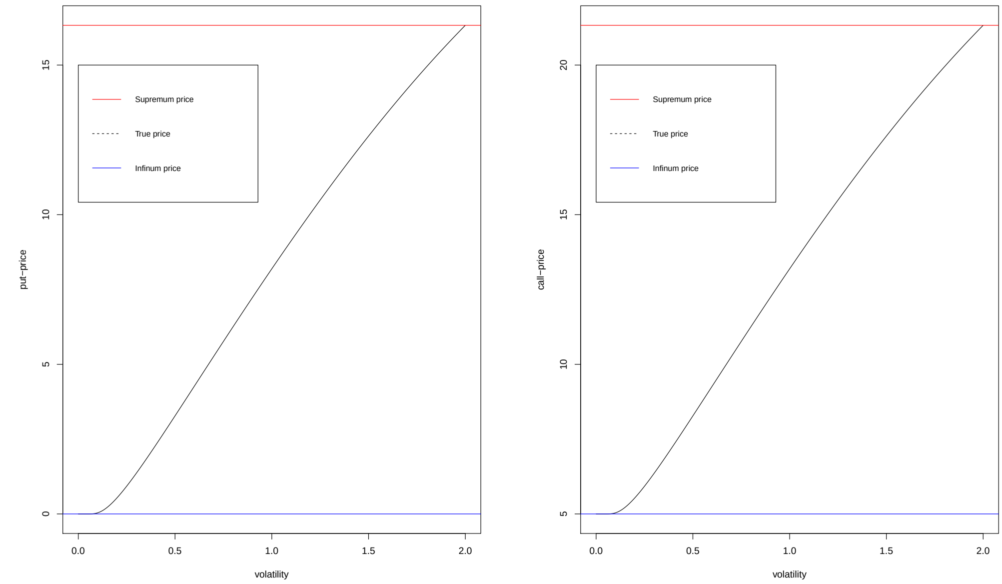
$$\sigma_i(t) = a_i \mathbb{1}_{[0,T_1]} + \sqrt{\frac{T\Sigma^2 - T_1 a_i^2}{T - T_1}}\mathbb{1}_{(T_1,T]}, \tag{131}$$

with $\Sigma < a_i < \Sigma\sqrt{T/T_1}$ for $i = 2, ...n$. Set $a_1 = \Sigma$ and let $\overline{a} = \max\{a_i\}$ and $\underline{a} = \min\{a_i\}$.
We now consider a the problem of a put option X with maturity $T_1 < T$. Then, for each i $(\Omega, \mathcal{F}, \{\mathcal{F}\}_t, Q_i)$ defines a complete market model. However, the delta of the Black & Scholes formula depends on the volatility structure, and is therefore not model free. Meaning that the delta-hedged position is almost surely zero with respect to Q_i, but not with respect to Q_j for $j \neq i$.

However, the Black & Scholes formula is monotonically increasing as a function of its volatility (see Proposition 14.4 in the appendix), hence $\overline{\pi}(X) = P^{BS}(t, \overline{a})$ and $\underline{\pi}(X) = P^{BS}(t, \underline{a})$. $\square$

Cont gives a similar, albeit shorter example in his paper regarding call options.

Numerically, we get that for call and put options with spot price $S(t) = 30$, strike price $K = 25$, interest rate $r = 0$ and maturity time $T = 1$ year, we find the prices in figure 2.



(a) Put option

(b) Call option

Figure 2: This figure shows plots of a call option and put option as functions of the volatility, with the worst case bid and ask prices in red and blue.

In the next example, we look at model uncertainty on Margrabe's formula

Example 11.22. Let $C^M(t,\sigma)$ denote the price of a Margrabe style spread option at time t. In the same way as in the Black & Scholes formula, we find that $\log(\phi(d_1)/\phi(d_2)) = \log(S_2(0)/S_1(0))$, and therefore we can in a similar manner to Proposition 14.4 show that Margrabe's formula is monotone and increasing as a function of σ, where $\sigma := \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}$. We may therefore repeat the argument in the example above, and find that $\overline{\pi}(X) = C^M(t,\overline{a})$ and $\underline{\pi}(X) = C^M(t,\underline{a})$. $\square$

We can also redo Example 11.22 numerically, and also include the worst case prices.

Example 11.23. If we consider the example above as a price where all parameters are given, except the correlation, the parameter that in this case faces uncertainty, we may construct the following numerical example: Let the stock prices be given as $S_1(t) = 80$, $S_2(t) = 90$, $\sigma_1 = 0.2$, $\sigma_2 = 0.4$, interest rate $r = 0$ and time to maturity $T = 1$ year, we find the evolution of the price as a function of the correlation as described in figure 3 below.

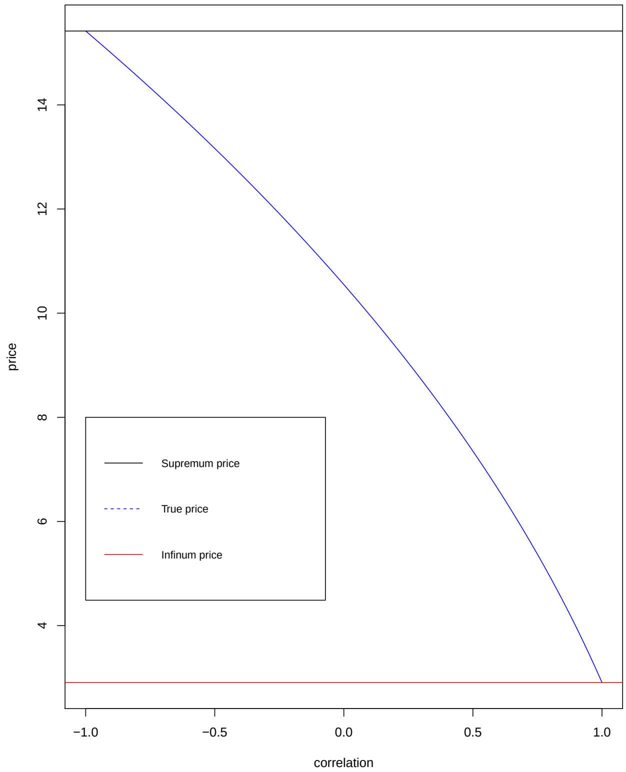


Figure 3: This figure shows the price of a spread option as a function of the correlation. The worst case scenario prices are plotted in as well. This figure shows clearly how big an impact uncertainty on the correlation parameter may have on the price of the option.

□

We refer to Cont [25] for more examples.

Why is this interesting? We have that the quantities $\bar{\pi}(X)$, $\underline{\pi}(X)$ and $\mu_{\mathcal{Q}}(X)$ can be used to compute a margin (for an over-the-counter instrument), or to provision for model uncertainty on the trade. Cont then shows that if the market price of a derivative X is computed using one of the pricing models $Q \in \mathcal{Q}$, for example Q_1, then the margin for the model uncertainty is

$$\bar{\pi}(X) - \mathbb{E}_{Q_1}[X] \leq \mu_{\mathcal{Q}}(X),$$

and in that way $\mu_{\mathcal{Q}}$ constitutes an upper bound on the margin for model risk. Furthermore, if $\pi_m(X)$ is the value of an option, then the *model risk ratio* (Cont equation (4.14)) defined as

$$MR(X) := \frac{\mu_{\mathcal{Q}}(X)}{\pi_m(X)}$$

can be used to indicate how large the component of the model risk is of the risk of a portfolio, and in that regard we can use the quantity $\mu_{\mathcal{Q}}$ for validating models. He summarizes the discussion of the usefulness by stating that the approach where we construct an interval $[\underline{\pi}(X), \bar{\pi}(X)]$ is much more compatible with the bid-ask spread, and therefore a much better measure of risk than other known risk measures, for example superhedging and maximal arbitrage bounds.

There are, however, problems with these kinds of measures for model risk, in the way that they require calibration with respect to some benchmark instruments, which may prove to be difficult when dealing with complex payoff structures. By relaxing the axioms of subadditivity and positive homogeneity and replacing them with convexity, we may introduce convex risk measures. As we will see, the calibration difficulties can be overcome by this change.

As noted earlier, any convex risk measure ρ has the representation

$$\rho(x) = \sup_{P \in \mathcal{P}} \{ \mathbb{E}[X] - \alpha(P) \}, \tag{132}$$

where $\alpha : \mathcal{P} \to \mathbb{R}$ is some penalty function, for which it is unclear what its financial interpretation should be. Moreover, equation (132) is not normalized, making it difficult to compare it to market valued portfolios.

The main idea in the section on convex measures of model risk in Cont's paper, is that the penalty function α is replaced by a *penalty function*, a function that penalizes each model by its pricing error $\|C^* - \mathbb{E}_Q[H]\|$ on the benchmark instruments. Also, we do not require the models to be calibrated to replicate the option prices on the benchmark instruments (or within a bid-ask spread). We define

$$\pi^*(X) := \sup_{Q \in \mathcal{Q}} \{ \mathbb{E}_Q[X] - \|C^* - \|C^* - \mathbb{E}_Q[X]\|\} \tag{133}$$

and

$$\pi_*(X) := \inf_{Q \in \mathcal{Q}} \{ \mathbb{E}_Q[X] + \|C^* - \mathbb{E}_Q[X]\|\}. \tag{134}$$

In the same way as with the coherent measure of model risk defined in (121), we may treat (133) and (134) in a similar way. That is, we define a mapping $\mu_* : \mathcal{C} \to \mathbb{R}$ by

$$\mu_*(X) := \pi^*(X) - \pi_*(X) \tag{135}$$

for all $X \in \mathcal{C}$.

We note that different choice of norm in (133) and (134) leads to different meaures in (135). For example, for any $p > 0$ we have

$$\|C^* - \mathbb{E}_Q[H]\|_p^p = \sum_{i \in I} |C_i^* - \mathbb{E}_Q[H_i]|^p,$$

and

$$\|C^* - \mathbb{E}_Q[H]\|_\infty = \sup_{i \in I} |C_i^* - \mathbb{E}_Q[H_i]|.$$

Then, we are able to state the main result in chapter 5 in Cont's paper.

Proposition 11.24. *Assume that the pricing error satisfies* $\|C^* - \mathbb{E}_Q[H]\| \geq |C_i^* - \mathbb{E}_Q[H_i]|$ *for all* $i \in I$ *and* $Q \in \mathcal{Q}$*, then*

1. π^* *assigns to any benchmark option a value lower than its market price. I.e*

   $$\pi^*(H_i) \leq C_i^*$$

   for all $i \in I$*. Furthermore, the converse is true for* π_**.*

2. *If* $\mathcal{Q}$ *contains at least one model such that* $\mathbb{E}_Q[H_i] = C_i^*$ *for all* $i \in I$*, then there is no uncertainty, meaning that*

   $$\pi_*(H_i) = \pi^*(H_i) = C_i^*$$

   for all $i \in I$*, and that for any payoff* $X \in \mathcal{C}$

   $$\pi^*(X) \geq \pi_*(X).$$

3. *If* $\mathcal{Q}$ *contains at least one model such that* $\mathbb{E}_Q[H_i] = C_i^*$ *for all* $i \in I$*, then* μ_* *is a measure of model uncertainty that satisfies all the properties outlined before.*

4. *If* $\mathcal{Q}$ *contains at least one model such that* $\mathbb{E}_Q[H_i] = C_i^*$ *for all* $i \in I$*, then diversifying a position using long positions in benchmark derivatives reduces model uncertainty*

   $$\mu_*\left(\lambda_0 X + \sum_{i=1}^{n} \lambda_i H_i\right) \leq \mu_*(X),$$

   provided that $\lambda_k > 0$ *for all* k *and* $\sum_i \lambda_i = 1$*.*

   In particular, any position that can be replicated by a convex combination of available derivatives has no model uncertainty, i.e if $X = \sum_k \lambda_k H_k$*, then* $\mu_*(X) \leq 0$*.*

Proof: We refer to the Appendix in Cont [25]. ■

Remark 11.25. The theorem above, is stated for markets where the bid price and market price coincide. As this usually is not the case, Cont proposes a new risk measure to accommodate for this. We get

$$\mu_*(X) = \sup_{Q \in \mathcal{Q}} \left\{ \mathbb{E}_Q[X] - \alpha_0(Q) \right\} - \inf_{Q \in \mathcal{Q}} \left\{ \mathbb{E}_Q[X] - a_0(Q) \right\},$$

where the convex penalty function a_0 is defined by

$$a_0(Q) = \sup_{i \in I} \left\{ \| \max\left(0,\ \mathbb{E}_Q[H_i] - C_i^{\text{bid}}, C_i^{\text{ask}} -\ \mathbb{E}_Q[H_i]\right) \| \right\}.$$

Remark 11.26. How can one use this in practice? One may, given a set of benchmark options choose a pricing model denoted $Q_1 \in \mathcal{Q}$ replicating the prices, which can be calibrated to the prices of the options. Cont suggests using one-dimensional diffusion models and the SABR model[28] used for equity and index derivatives and European options on interest rates respectively, which are (according to Cont) easy to calibrate, but may suffer from being unrealistic when it comes to future scenarios. One may then add the calibrated model Q_1 to a sequence of other pricing models $\{Q_k\}_{k=2}^n$ whose features are more realistic, but may be difficult to calibrate. With the approach outlined above, for $k \geq 2$, we do not need to calibrate Q_k, but instead penalize the pricing error. Then Q_1 anchors μ_* in the set of market prices, and one may then incorporate the Q_k without heavy numerical procedures. Moreover, The constraint of including one pricing model such that $\mathbb{E}_Q[H_i] = C_i^*$ ensures that the market is arbitrage-free, which is the same s requiring that $\rho(0) = 0$.

11.3 Risk capturing functionals

The notion of using convex risk measures as defined above by Cont and many others, has been further generalized by Bannör & Scherer [5], to what they call *risk capturing functionals*. They introduce it mainly as a tool to quantify parameter risk, but also in the more general setting of model risk. They state a list of three properties that such a functional denoted Γ should satisfy:

[28]The SABR model describes a single forward F and its associated volatility σ, whose dynamics are given by

$$\begin{aligned} \mathrm{d}F(t) &= \sigma_t (F(t))^\beta \, \mathrm{d}W(t) \\ \mathrm{d}\sigma_t &= \alpha \sigma_t \, \mathrm{d}Z(t), \end{aligned}$$

where $\beta \in [0, 1]$, $\alpha \geq 0$ and W and Z are two correlated Wiener processes.

Scherer.1 Order preservation: If there exists a model-free order, it should be preserved when incorporating model uncertainty. That is, for contingent claims X, Y:

$$X(\omega) \geq Y(\omega) \text{ for all } \omega \in \Omega \implies \Gamma(X) \geq \Gamma(Y).$$

Scherer.2 Diversification: Diversification of model uncertainty should not be penalized, meaning that a convex combination of two positions facing model uncertainty should not have a higher price than a convex combination of the individual prices. Hence

$$\Gamma\left(\lambda X + (1-\lambda)Y\right) \leq \lambda\Gamma(X) + (1-\lambda)\Gamma(Y)$$

must hold for any contingent claim X and for all $0 \leq \lambda \leq 1$.

Scherer.3 Model independence consistency: If a contingent claim X is consistently priced under all models $Q \in \mathcal{Q}$, then no model uncertainty is present and the model risk-captured price agrees with the risk-neutral price. Hence no charge for model risk is added to the risk-neutral price. Meaning that

$$Q \mapsto \mathbb{E}_Q[X] \text{ is constant on } Q \implies \Gamma(X) = \mathbb{E}_Q[X].$$

We can now define the model risk capturing functionals as proposed by Bannör & Scherer. The idea is basically that the model risk of a derivative X can by measured by applying a convex risk measure to the evaluation mapping.

Definition 11.27 (Model risk capturing functional)**.** Let $\mathcal{Q}$ be a family of models and let R be a probability measure on $\mathcal{Q}$. Let $\mathcal{A} \subset \mathcal{L}^0(R)$ be a vector space of measurable functions containing the constants. Then denote

$$\mathcal{C}^{\mathcal{A}} := \left\{ X \in \bigcap_{Q \in \mathcal{Q}} L^1(Q) : \epsilon_X : Q \mapsto \mathbb{E}_Q[X] \in \mathcal{A} \right\}$$

as the vector space of all $\mathcal{A}$-regular claims being available for all models in the model family $\mathcal{Q}$. Furthermore, let $\rho : \mathcal{A} \to \mathbb{R}$ be a normalized, law invariant convex risk measure. Then the mapping $\Gamma : \mathcal{C}^{\mathcal{A}} \to \mathbb{R}$ defined by

$$\Gamma(A) := \rho(\epsilon_X) = \rho(Q \mapsto \mathbb{E}_Q[X]) \tag{136}$$

is called a *model risk capturing functional* on the set of claims $\mathcal{C}^{\mathcal{A}}$ w.r.t. the distribution R. ρ is called the *generator of* Γ and $\Gamma(A)$ is called the *risk captured ask price of* X given the functional Γ.

They state that canonical choices for $\mathcal{A}$ are the $L^p(R)$ spaces for $p \in [1, \infty]$.

The idea behind using these tools, are according to Bannör & Sherer, that a trader facing model risk should choose high enough ask prices and low enough bid prices that the trader has some kind of buffer to prevent losses due to model risk. Also, when buying derivatives, to account for model risk, one should set the bid prices low enough to prevent losses when choosing a model.

One of the more popular convex risk measures, is the *Average-Value-at-Risk*, or *AVar* for short. Value-at-Risk (VaR) is, according to Hull [44], *"an attempt to provide a single number summarizing the total risk in a portfolio of financial assets"*, and we refer to his book for a precise definition. However, it is known that VaR is not a coherent or convex risk measure (see e.g Dahl [26]). AVaR, on the other hand, is a coherent law-invariant risk measure (see e.g Föllmer & Schied [40]). AVar is defined as

$$\mathrm{AVaR}_\alpha(X) := \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_\beta(X)\,\mathrm{d}\beta, \tag{137}$$

where X is integrable and $0 < \alpha \leq 1$.

Using this, Bannör & Scherer introduce the Average-Value-at-Risk-induced risk capturing functional.

Definition 11.28. Let $\mathcal{Q}$ be a family of martingale measures inducing model risk, and let R be a distribution on $\mathcal{Q}$. Consider the $L^1(R)$-regular claims then define the Average-Value-at-Risk-induced risk capturing functional $R * \mathrm{AVaR}_\alpha : \mathcal{C}^{L^1(R)} \to \mathbb{R}$ defined as

$$R * \mathrm{AVaR}_\alpha(X) := AVaR_\alpha(\mathcal{Q} \mapsto \mathbb{E}_{\mathcal{Q}}[X]). \tag{138}$$

$R * \mathrm{AVaR}_\alpha$ is then the risk capturing functional generated by the coherent risk measure AVaR_α for a given confidence level α as defined above.

The functional $R * \mathrm{AVaR}_\alpha$ captures the model risk quantified by the distribution R, and deals with the risk from the upper α-tail of the price distribution with respect to different parameters by averaging over the tail prices. This is in contrast to the regular Average-Value-at-Risk, which captures the risk within a specific model. The Average-Value-at-Risk-induced risk capturing functionals generalizes the framework of Cont (see his 2006 paper [25] or earlier chapters in this thesis) in the following way. If we choose a distribution R on $\mathcal{Q}$, the essential supremum becomes the regular supremum, and therefore

$$R * \mathrm{AVaR}_0(X) = \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q[X] = \bar{\pi}(X).$$

holds, implying that if we have uncertainty on at most countably many models, Cont's worst case framework agrees with Bannör & Scherer's Average-Value-at-Risk-induced risk capturing functional in the sense that Cont uses the extreme points of $R{*}\mathrm{AVaR}$, and the method does then allow us to interpolate between these methods, and they therefore state that $R{*}\mathrm{AVaR}$ provides prices with extra charge for model uncertainty, but being more conservative than expected values and less conservative than a supremum. Furthermore, Bannör & Scherer then introduce the notion of entropic-induced risk-capturing functionals, which are valid for claims $X \in L^{\infty}(P)$. This is an alternative generalization of Cont's framework, which may be used to account for risk associated with large trades which may bear risk due to liquidity effects or other factors. This is due to the fact that the entropic risk measure is not positive homogeneous. We refer to Bannör & Scherer [5] for details. They also provide several results on the convergence and asymptotics of the risk-captured prices.

We end this section with a brief review of a case-study, done on a European Margrabe type exchange option, which is found in chapter 5.1 in Bannör & Scherer [5]. The set-up is for the most part the same as in chapter 5.1 in this thesis, but with the addition of a non-zero interest rate and drift term in the dynamics of the stocks. None of these things matter, the end result becomes the same anyway.[29] Then, they assume all parameters are known, except for the correlation ρ, which they estimate via Pearson's correlation coefficient, defined in equation (94)[30]. Then they apply the Fisher transform [36, 37] as we did in section 10.1, and they find that for $n \gg 3$

$$\operatorname{arctanh}(\rho) \sim N\left(\operatorname{arctanh}(\rho_0), \frac{1}{n-3}\right),$$

in order to numerically calculate the prices. Above, ρ_0 denotes the true correlation.They then apply the risk-capturing functionals and find the following

[29] This is of course only true for situations of the Margrabe kind, and *not* when we are dealing with forward contracts and tap into the realm of Black-76.

[30] Note that is missing two $\sqrt{n}$ in the denominator in the estimator for ρ. The estimator in (94) is the correct.

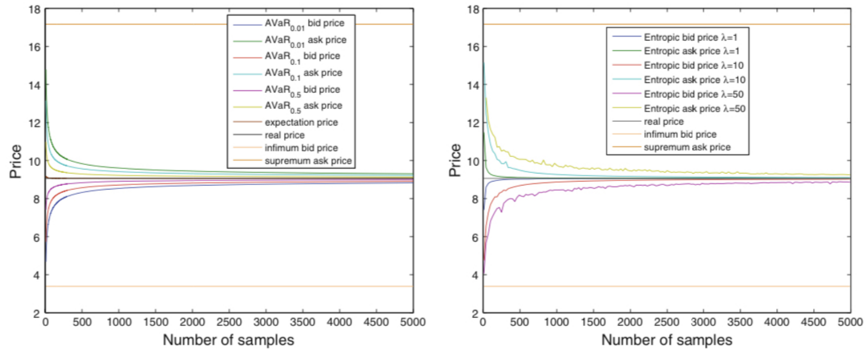


Figure 4: This figure shows AVaR-induced bid-ask spreads and entropic-induced bid-ask spreads together with the real price, expected price and the worst case bid-ask spreads for different significance levels. This figure is **Fig. 3** in Bannör & Scherers's article *Capturing parameter risk with convex risk measures*, European Actuarial Journal, 2013, pp 97-132, Springer Nature. Used with permission from Springer.

We see in figure 4, that on the left, they have plotted the AVaR-induced bid-ask spreads for different significance levels. We see that a higher significance level indicates a larger bid-ask spread. On the right they have done the same but with entropic-induced bid-ask prices. And again, a higher risk aversion parameter λ, leads to a bigger bid-ask spread. We also see that our findings about Cont's worst case approach to the bid-ask spread is much wider than other approaches.

11.4 Extension to other spaces

The notion of coherent and convex risk measures, as defined by Artzner et al. [2] and Föllmer & Schied [40] among others, is usually only defined for random variables $X \in L^\infty$. That is, the space of essentially bounded random variables. However, there is an increasing number of articles being written about the extensions to other spaces, most notably L^p, the space of random variables with finite p-moment. Notable examples of authors are Filipović & Svindland [35], Kaina & Rüschendorf [47] and Ruszczyński & Shapiro [70, 69]. Filipović & Svindland show by example how to construct the risk measures Average-Value-at-Risk and entropic risk measures and L^p, as well as an L^p-extension of the worst case risk measure and several other examples.

In a similar fashion, we will define the theory of risk measures on the Filipović space H_w.

12 Risk measures on infinite dimensional spaces

We will now have a look at how we can expand the idea of risk measures to infinite dimensional spaces. We start by some background materials on risk measures where the random variables takes values in $\mathbb{R}$.

12.1 Background materials and a representation result

We start by introducing some notation and some definitions.

Definition 12.1 (Notation)**.** In this chapter, $(\Omega, \mathcal{F})$ denotes a measurable space, $\mathbb{V}$ denotes the set of all measures on $(\Omega, \mathcal{F})$ and $\mathbb{X}$ denotes a linear space of random variables $X : \Omega \to \mathbb{R}$ in which the constant functions are contained.

Then, in the same way as in the finite dimensional case, we may define acceptable random variables and the acceptance set in terms of the function ρ.

Definition 12.2 (Acceptable random variable)**.** A random variable X is said to be *acceptable* if $\rho(X) \leq 0$. X is said to *unacceptable* if $\rho(X) > 0$. We may therefore define the *acceptance set* $\mathcal{A}_\rho$ as

$$\mathcal{A}_\rho = \{X \in \mathbb{X} : \rho(X) \leq 0\}. \tag{139}$$

We then introduce some topological conditions, which needs to be fulfilled in order to get the representations from theorems 11.14 and 11.15. They are, paired spaces, compatible topologies and pairings as a bilinear form between two spaces.

Definition 12.3 (Pairing)**.** Let Y and V be two linear spaces. A pairing of Y and V is a bilinear form $\langle \cdot, \cdot \rangle$ on $Y \times V$, where for all $y \in Y - \{0\}$ there exists some $v \in V$ such that $\langle y, v \rangle \neq 0$.

Definition 12.4 (Compatible topology)**.** A topology on Y is compatible with the pairing if it is a locally convex topology such that the linear function $\langle \cdot, v \rangle$ is continuous for all $v \in V$. Moreover, any continuous linear function on Y can be written on this form for some $v \in V$.

A compatible topology on V is defined in a similar fashion.

Definition 12.5 (Paired spaces)**.** Two linear spaces Y and V are said to be *paired spaces* if there exists a pairing between Y and V and the two spaces have compatible topologies with respect to the pairing.

We illustrate the point of paired spaces with an example on L^p spaces, where we show that L^p and L^q are paired spaces proved that $1/p + 1/q = 1$.

Example 12.6. Any vector space V together with its dual space V^* and the bilinear map $\langle x, f\rangle := f(x)$, where $x \in V$ and $f \in V^*$ form a dual pair. Consider for example the L^p spaces, and define p and q such that $\frac{1}{p} + \frac{1}{q} = 1$. Let $f \in L^p$ then we know from Riesz Representation Theorem (see e.g McDonald & Weiss [55] **Theorem 13.12**) that there exists a unique $g \in L^q$ such that

$$l(f) = \langle f, g\rangle = \int_A fg \, \mathrm{d}\mu$$

and $\|l\|_{p^*} = \|g\|_q$. Therefore, whenever $\frac{1}{p}+\frac{1}{q} = 1$, L^p and L^q are paired spaces.

We end this section with three definitions that we need to impose on the function ρ in order to get the representation in theorems 11.14 and 11.15. Then a final definition on the convex conjugate of a function.

Definition 12.7 (Convex function)**.** Let $C \subset Y$ be a convex set, and let $x, y \in C$ and $\lambda \in [0, 1]$. A function $f : C \to \mathbb{R}$ is a *convex function* if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y). \tag{140}$$

Definition 12.8 (Proper function)**.** Let A be a subset of a linear space X, and let $f : A \to \mathbb{R}^*$ be a function. f is a *proper function* if $\mathrm{dom}(f) \neq \emptyset$ and $f(x) > -\infty$ for all $x \in A$.

Definition 12.9 (Lower semicontinuous function)**.** Let $A \subset Y$ be a set, and let $f : A \to \mathbb{R}^*$ be a function. f is *lower semicontinuous* (often abbreviated as lsc) if for each $r \in \mathbb{R}$ such that $r < f(x_0)$ there exists a neighborhood U of x_0 such that $f(u) > r$ for all $u \in U$. Equivalently, f is lower semicontinuous if

$$\liminf_{x \to x_0} f(x) \geq f(x_0).$$

Definition 12.10 (Convex conjugate)**.** Let Y and V be paired spaces, and let $f : Y \to \mathbb{R}^*$ be a function. Then the *convex conjugate* $f^* : V \to \mathbb{R}^*$ is defined by

$$f^*(v) = \sup\left\{\langle y, v\rangle - f(y) : y \in Y\right\}.$$

The *biconjugate* $f^{**} : Y \to \mathbb{R}^*$ is defined as

$$f^{**}(y) = \sup\left\{\langle y, v\rangle - f^*(y) : v \in V\right\}.$$

Theorem 12.11 (Infinite representation of risk measures)**.** *Let $\mathbb{X}$ be a vector space paired with the linear space V as defined above. Let $\rho : \mathbb{X} \to \mathbb{R}$ be a proper, lower semicontinuous and convex function. We have from **Theorem 1.7.9** in Dahl [26] that*

$$\rho(X) = \sup\left\{\langle X, v\rangle - \rho^*(v) : v \in \mathrm{dom}(\rho^*)\right\}.$$

Then

1. *ρ is monotone if and only if every $v \in \operatorname{dom}(\rho^*)$ are such that $v \leq 0$.*
2. *ρ is translation invariant if and only if $v(\Omega) = -1$ for all $v \in \operatorname{dom}(\rho^*)$.*

Hence if ρ is a convex risk measure (i.e monotonicity and translation invariance holds), then $v \in \operatorname{dom}(\rho^*)$ implies that $Q := -v \in \mathcal{P}$ and

$$\begin{aligned}\rho(X) &= \sup_{Q \in \mathcal{P}} \{\langle X, -Q \rangle - \rho^*(-Q)\} \\ &= \sup_{Q \in \mathcal{P}} \{\langle -X, Q \rangle - \rho^*(-Q)\} \\ &= \sup_{Q \in \mathcal{P}} \left\{ \mathbb{E}_Q[-X] - \alpha(Q) \right\},\end{aligned}$$

where $\alpha(Q) = \rho^*(-Q)$ is a penalty function, and the pairing $\langle \cdot, \cdot \rangle$ is viewed as an expectation.

We end this example with an example, where $\mathbb{X}$ is an L^p space, and allows us to revisit the previous example and illustrate Theorem 12.11. The extension $\mathbb{X} = L^p(\Omega, \mathcal{A}, P)$ is briefly discussed in the next chapter. This example is loosely based on Dahl [26] and Filipović & Svindland [35].

Example 12.12. Let $\mathbb{X} = L^p(\Omega, \mathcal{F}, P)$, and let $\mathbb{V}$ be the vector space of all signed measures on $L^p(P)$, and let $V \subset \mathbb{V}$ be the set of signed measures v such that for all $X \in L^p(P)$

$$\int_\Omega |X| \, \mathrm{d}|v| < \infty.$$

By the Jordan decomposition theorem, we can find a decomposition of v such that $v = v_p - v'$, where $v_p \ll P$.

We want to construct a pairing of $L^p(P)$ and $(L^p(P))^*$. We claim that this pairing is given by $(X, Z) \mapsto \mathbb{E}[XZ]$, for $X \in L^p(P)$ and $Z \in (L^p(P))^*$. We know from **Theorem 11, p. 79** in Lax [49] that $(L^p(P))^* = L^q(P)$, where $q = \frac{p}{p-1}$, or equivalently $1/p + 1/q = 1$. See the footnote[31] for a proof. Clearly this quantity is bilinear, so it suffices to check that it is well defined. Since $v_p \ll P$, we have by the Radon-Nikodym theorem that

$$v_p(A) = \int_A f \, \mathrm{d}P,$$

for all $A \in \mathcal{F}$. We then need to check that $f \in L^q(P)$. We get that $\int_A |f|^q \, \mathrm{d}P = \int_A |f|^{q-1} \, \mathrm{d}|v_p| < \infty$, so we get that any element in $L^q(P)$ may be represented

[31] $1/p + 1/q = 1$ implies that $1/q = (p-1)/p$. A simple rearranging yields the desired result.

by the Radon-Nikodym derivative with respect to P. Therefore, for $X \in L^p(\Omega)$ and $Z \in L^q(\Omega)$ we find that

$$\langle X, Z\rangle = \int_A XZ \,\mathrm{d}P = \int_A Xf \,\mathrm{d}P,$$

where we in the last equality expressed $Z \in L^q(\Omega)$ by the Radon-Nikodym derivative as described earlier. By substituting $\mathrm{d}P = \frac{\mathrm{d}v_p}{f}$, we find that

$$\int_A Xf \,\mathrm{d}P = \int_A X \,\mathrm{d}v_p \leq \int_A |X| \,\mathrm{d}|v_p| < \infty,$$

by the assumptions on the space $\bar{V}$. We therefore find that $\mathbb{E}[XZ]$ is a well-defined pairing of $L^p(P)$ and $V = L^q(\Omega)$. And therefore, a convex risk measure on $L^p(P)$ has the representation

$$\sup_{P \in \mathcal{P}} \left\{ \int_\Omega -X \,\mathrm{d}v_p - \rho^*(P) \right\}.$$

12.2 Risk measures on the Filipović spaces F^p

In this section, our aim is to define risk measures on the Filipović space, as described in Filipović [34] and Benth & Krühner [11, 12]. As we have seen in the section about infinite dimensional pricing of financial derivatives, we evaluate a function $g \in H_w$ at a point, and the pricing algorithms gives us a number in $\mathbb{R}$. It will therefore make sense to define risk measures on this space. In order to pursue generality, we expand the theory about H_w to a more general class of function spaces, which I have decided to call the Filipović spaces, denoted as $F^p(\Omega, \mathcal{A}, \mu)$, where H_w is a special case.

12.2.1 (†) The construction the Filipović spaces F^p

In this section, we will expand the Filipović space and define risk measures on this expansion.

Definition 12.13. Let $p \in (0, \infty)$. The set of $H_{w,p}$-integrable functions is defined to be

$$H_{w,p}(\mathbb{R}, \mathcal{A}, \mu) = \Big\{ f : \mathbb{R}_+ \to \mathbb{R}_+ : f \text{ is absolutely continuous and} \\ \int_{\mathbb{R}_+} |f'|^p w \,\mathrm{d}\mu < \infty \Big\}.$$

We define the norm as

$$\|f\|_{w,p}^p = |f(0)|^p + \left(\int_{\mathbb{R}_+} |f'|^p w \,\mathrm{d}\mu \right).$$

We note that the w, p-integrable functions are exactly those functions for which $\|f\|_{w,p} < \infty$.

Definition 12.14. Two functions $f, g \in H_{w,p}$ are equivalent if $f = g$ μ-almost everywhere. This then becomes

$$f \sim g \iff f = g \ \mu - \text{almost everywhere.}$$

We can easily prove that $\sim$ above is an equivalence relation.

Lemma 12.15. *It is clear that $\sim$ is reflexive and symmetric. The transitivity then follows since*

$$\|f - h\|_{w,p} = \|f - g + g - h\|_{w,p} \leq \|f - g\|_{w,h} + \|g - h\|_{w,p} = 0,$$

which concludes the proof.

We denote the equivalence class of a function $f \in H_{w,p}(\mathbb{R}_+, \mathcal{A}, \mu)$ by $[f]$, similarly as the way the L^p-spaces are defined.

We can now defined the Filipović spaces, denoted $F^p(\Omega, \mathcal{A}, \mu)$.

Definition 12.16 (The Filipović spaces F^p)**.** Let $1 \leq p < \infty$, and let $\sim$ be the equivalence relation defined as above. We define $F^p(\Omega, \mathcal{A}, \mu)$ to be the quotient space $H_{w,p}/\sim$. Meaning that

$$F^p(\Omega, \mathcal{A}, \mu) = \{[f] : f \in H_{w,p}(\Omega, \mathcal{A}, \mu)\}$$

We see that the Filipović spaces are linear spaces. We can now prove that F^p is a separable Banach space for all $1 \leq p < \infty$.

Theorem 12.17 (†)**.** *F^p is a separable Banach space for all $1 \leq p < \infty$.*

Proof: It is known that a product of Banach spaces is again a Banach space, and a product of separable spaces is again separable. Therefore, the space $\mathbb{R}_+ \times L^p(\mathbb{R}_+)$ equipped with the norm $\left(|\cdot|^p + \|\cdot\|^p_{L^p(\mathbb{R}_+)}\right)^{1/p}$ is a separable Banach space. We will show that F^p is isometrically isomorphic to $\mathbb{R}_+ \times L^p(\mathbb{R}_+)$. Similarly to the proof of **Theorem 5.1.1** in Filipović [34], define the linear operator $T : F^p \to \mathbb{R}_+ \times L^p$ by

$$Tf = \left(f(0), f' w^{1/p}\right).$$

We claim that T is an isometrical isomorphism. We start by proving that T is an isometry. From the Filipović -norm we find that

$$\|f - g\|^p_{w,p} = |f(0) - g(0)|^p + \int_{\mathbb{R}_+} |f'(x) + g'(x)|^p w(x)\, \mathrm{d}x.$$

In a similar fashion, we find that

$$\begin{aligned}\|Tf - Tg\|_{\mathbb{R}\times L^p(\mathbb{R}_+)} &= \|T(f-g)\|_{\mathbb{R}_+\times L^p(\mathbb{R}_+)}\\ &= |f(0)-g(0)|^p + \int_{\mathbb{R}_+} |f'(x)+g'(x)|^p w(x)\,\mathrm{d}x\\ &= \|f-g\|_{w,p}^p,\end{aligned}$$

which proves that T is an isometry. We note that $\|x-y\|$ is the metric induced by the norms of the respective spaces.

In order for the linear operator T to become an isometric isomorphism, we have to prove that it is a continuous bijection. The continuity criterion can be proven in two ways. The easiest method is just by observing that each component of T is continuous by definition. The other follows by knowing that *all* isometries $S : X \to Y$ are continuous. The proof is seen by simply choosing $\delta = \epsilon$ and then $d_X(x,y) = d_Y(f(x), f(y))$ we have that if $d_X(x,y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$.

We will now prove that T is a bijection. We do this by proving the existence of its inverse operator T^{-1}. We claim that the operator $T^{-1} : \mathbb{R}_+ \times L^2(\mathbb{R}_+)$ given by

$$T^{-1}(x,f)(u) = x + \int_0^u f'(\eta)w^{-1/p}(\eta)\,\mathrm{d}\eta$$

is the inverse operator of T. We see that

$$\begin{aligned}T^{-1}Tf(x) &= T^{-1}\left(\left(f(0), f'w^{1/p}\right)\right)\\ &= f(0) + \int_0^x f'(\eta)w^{1/p}(\eta)w^{-1/p}(\eta)\,\mathrm{d}\eta\\ &= f(0) + \int_0^x f'(\eta)\,\mathrm{d}\eta\\ &= f(x),\end{aligned}$$

where the least equality follows from Lemma 8.34, since it is exactly the definition of the weak derivative. We therefore find that $T^{-1}T = \mathrm{Id}$, which proves that T is a bijection, and therefore an isometric isomorphism. It therefore follows that for all $1 \le p < \infty$ $F^p(\mathbb{R}_+)$ is isometrically isomorphic to $\mathbb{R}_+ \times L^p(\mathbb{R}_+)$, from which it follows that they are separable Banach spaces. ■

From this it follows as a corollary that $F^2(\mathbb{R}_+)$ is a separable Hilbert space. This follows immediately from Theorem 12.17, or it can be proven directly, as done by Filipović in **Theorem 5.1.1** in Filipović [34].

12.2.2 Risk measures on F^p

As we have seen in earlier sections, the infinite representations of the contingent claims have all taken in a function g from the Filipović space H_w, and then after evaluating at a point x, we get a real number. We therefore claim that all these representations live in the space $F^2(\mathbb{R}_+)$. We therefore define as in the previous section $\mathbb{X} = F^p(\mathbb{R}_+)$. That is, $\mathbb{X}$ is the set of functions $X : F^p(\mathbb{R}_+) \to \mathbb{R}$. We will therefore define the convex risk measures on the Filipović spaces in the following manner:

Definition 12.18. A function $\rho : F^p(\mathbb{R}_+) \to \mathbb{R} \cup \{\infty\}$ is a *convex risk measure* on $F^p(\mathbb{R}_+)$ if

(i) $\rho(0) < \infty$

(ii) The measure ρ is convex, meaning that for all $0 \leq \lambda \leq 1$ and $X, Y \in F^p(\mathbb{R}_+)$
$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$$

(iii) The measure ρ is cash-invariant], meaning that for $m \in \mathbb{R}$ and $\mathbf{1} \in F^p(\mathbb{R}_+)$, where $\mathbf{1}$ is the function that is constantly 1
$$\rho(X + m\mathbf{1}) = \rho(X) - m$$

(iv) The measure ρ is monotone in the sense that if $X, Y \in F^p(\mathbb{R}_+)$ and $X \geq Y$, then $\rho(X) \leq \rho(Y)$.

Remark 12.19. We note that in the same way as earlier, that if ρ is positive homogeneous in the sense that $\rho(tX) = t\rho(X)$ for all $t \geq 0$, then ρ is a *coherent risk measure* on $F^p(\mathbb{R}_+)$.

We may also define the *acceptance set* associated with ρ:

Definition 12.20. The set of acceptable risky positions is denoted $\mathcal{A}_\rho$, and is defined as
$$\mathcal{A}_\rho = \{X \in F^p(\mathbb{R}_+) : \rho(X) \leq 0\}.$$

We therefore see that this definition is more or less the same as the acceptance set defined in (116). And likewise, if $\mathcal{A} \subset F^p(\mathbb{R}_+)$ is a set of acceptable random variables (that is, the members of $\mathcal{A}_\rho$) then the set $\mathcal{A}$ has an associated risk measure $\rho_{\mathcal{A}} : F^p(\mathbb{R}_+) \to \mathbb{R} \cup \{\infty\}$ defined as
$$\rho_{\mathcal{A}} = \inf\left\{m \in \mathbb{R} : m + X \in \mathcal{A}_\rho\right\}.$$

Theorem 11.13 about the properties of ρ as in the earlier case does not assume anything about the space on which the measures operate, and therefore

Theorem 12.21. *Let ρ be a convex risk measure with induced acceptance set $\mathcal{A}_\rho$. Then*

(i) $\rho_{\mathcal{A}_\rho} = \rho$

(ii) $\mathcal{A}_\rho$ *is a nonempty convex set*

(iii) *If* $X \in \mathcal{A}_\rho$ *and there exists* $Y \in F^p(\mathbb{R}_+)$ *such that* Y *dominates* X*, then* $Y \in \mathcal{A}_\rho$

(iv) ρ *is a coherent risk measure only if* $\mathcal{A}_\rho$ *is a convex cone*

Conversely, let $\mathcal{A}$ *be a nonempty convex subset of* H_w*. Let* $\mathcal{A}$ *be such that if* $X \in \mathcal{A}$ *and* $Y \in F^p(\mathbb{R}_+)$ *such that* Y *dominates* X *implying* $Y \in \mathcal{A}$*, then the following holds true*

(v) $\rho_{\mathcal{A}}$ *is a convex risk measure*

(vi) $\mathcal{A}$ *is a convex cone only if* $\rho_{\mathcal{A}}$ *is a coherent risk measure*

(vii) $\mathcal{A} \subset \mathcal{A}_{\rho_{\mathcal{A}}}$.

Moreover, if we let $\mathbb{V}$ denote the linear space of all finite signed measures on $F^p(\mathbb{R}_+)$, and let $V \subset \mathbb{V}$ be defined to set of $v \in \mathbb{V}$ such that $\int_{\mathbb{R}_+} |X| \, \mathrm{d}|v| < \infty$. If we let $F^p(\mathbb{R}_+)$ and V have topologies such that they become paired spaces. We have shown that all normed spaces carry a locally convex topology, so we assume that this topology makes the pairing $\langle \cdot, v \rangle$ continuous. We therefore get from **Theorem 1.7.9** in Dahl [26] that $\rho(X) = \sup\{\langle X, v\rangle - \rho^*(v) : v \in \operatorname{dom}(\rho^*)\}$. Then, assuming that ρ is convex, lower semicontinuous and proper, we can represent ρ as

$$\begin{aligned}
\rho(X) &= \sup_{Q \in \mathcal{P}} \{\langle X, -Q\rangle - \rho^*(-Q)\} \\
&= \sup_{Q \in \mathcal{P}} \{ \mathbb{E}_Q[-X] - \alpha(Q)\},
\end{aligned}$$

where $\alpha(Q) = \rho^*(-Q)$ as in Theorem 12.11 and the pairing $\langle \cdot, \cdot \rangle$ is viewed as an expectation.

13 Discussion

13.1 How option prices depend on the correlation/covariance

In the classical Black & Scholes model where one wants to price an option on a single asset, the biggest concern is the volatility - being the only quantity not given directly by the market. However, there exists many methods to estimate the historical volatility, for example by annualizing the standard deviation of daily log-returns, or by reverse engineering by using option prices from the market and then using that information to find the volatility. The latter approach is somewhat useless for more exotic options, since they often are sold "over the counter" and are in general not available (or liquid enough) to be seen directly in the market. It has been shown by Lindström [51] that if we were able to observe the market filtration $\mathcal{F}_t$, then we would be able to estimate the volatility without errors, which in turn shows that the Black-Scholes model does not have any internal inconsistencies using given $\mathcal{F}_t$. The market filtration is, however, not available to the market participants, and using the *observable filtration*, some internal inconsistencies can be found.

We will in the next example show the first method of volatility estimation, on market data from the S&P500 index.

Example 13.1. The market has by convention 252 trading days, so using this going back one year from today's date (August 16, 2019) we go back to August 17, 2018. The logreturns are defined as $x(t) = \log(S(i+1)) - \log(S(i))$, where $S(i)$ denotes the share price on the i'th day. We then find that $\sigma = 16.18\%$ annually. And since the closing price was \$2,847.60, we may appeal to the Black & Scholes formula [14] to find the price for plain vanilla options on this index for any strike price. $\square$

With this in mind, we will now have a look at the easiest (at least in a market sense) spread option, whose prices are given by Margrabe's formula [53]. It is known that the Black & Scholes formula is monotonically increasing as a function of its volatility (see Appendix 2 for a quick proof). We have seen earlier that if we define $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}$, then Margrabe's formula possesses the same property.

Example 13.2. Let $s_1 = 10$ and $s_2 = 12$ in your favorite currency. Let the volatilities be constant and be set to $\sigma_1 = 10\%$ and $\sigma_2 = 20\%$, and let the spread option have a strike time of half a year. Then, if we want to price the option as a function of the correlation, we find that

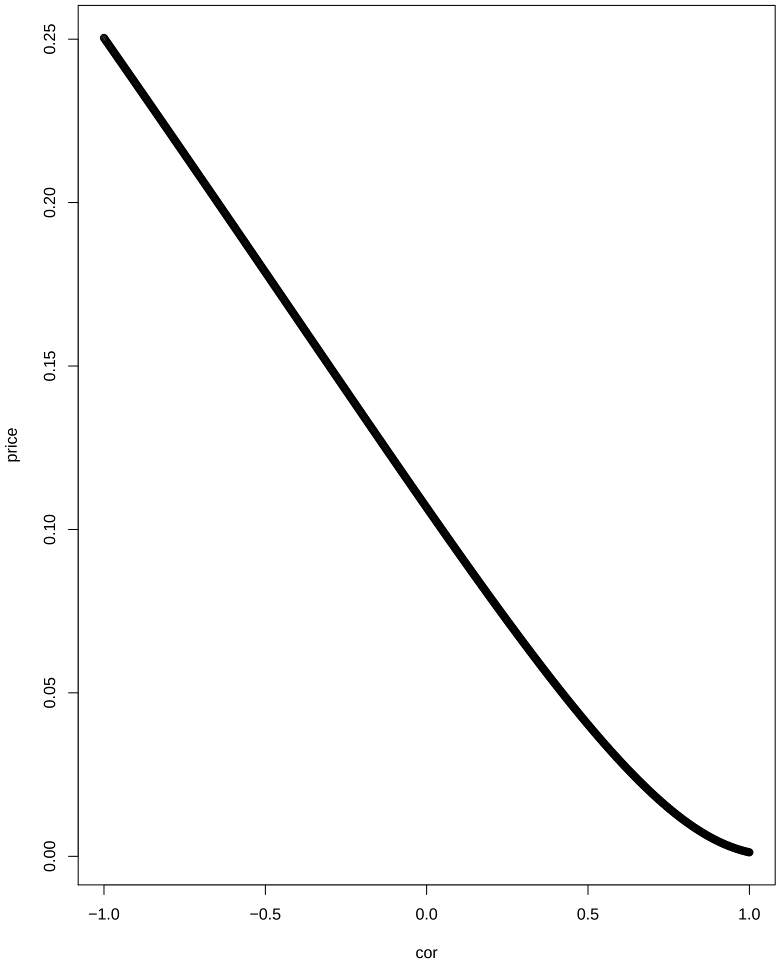


Figure 5: Plot showing the price of a simple spread option as a function of the correlation ρ.

13.2 Conclusion and further studies

We have in this thesis shown that in the same way as in the n-dimensional case, the option prices depend on the volatility structure, or to be more precise, the covariance structure of the underlying. And as mentioned earlier, since the true convariance operator is not given to us, we have provided results allowing market practitioners to estimate this operator. We have also given a thorough discussion on the notion of model risk, with emphasis on convex risk measures and risk capturing functionals. Since the Filipović space has proven itself to be quite useful in the stochastic modelling of electricity futures (and related), one object of further study would be to properly and rigorously define and prove more results on risk measures on the Filipović space, much in the same way as Filipović & Svindland [35]. It would also be interesting to further develop the notion of risk capturing functionals, and extending the work already done on this topic to the Filipović space, or more generally separable Hilbert spaces (or even Banach spaces). Another example of an interesting result, would be to prove, like in the n-dimensional case, results on the monotonicity of the option prices in Hilbert spaces with respect to the covariance structure. I have been working on this, but to my knowledge it requires the Gateaux derivative, which is beyond the scope of this thesis (which is already long enough).

However, some other projects springs to mind.

13.3 Stochastic and rough volatility for infinite dimensional representations of commodity pricing

One thing I saw during the literature hunt for my thesis, was that all of the literature I found on the subject of infinite dimensional pricing and hedging, assumed that the volatility of the driving noise should be modelled as a deterministic function. There is already some literature on this when it comes to interest rates and the theory of fixed income markets, but to my knowledge none when it comes to commodity markets. One interesting area of research, would be to find conditions under which we can state and prove "nice" closed form formulas for infinite dimensional contingent claims, like in the case where the volatility is deterministic. Or, if that is not possible, prove (or at least substantiate) that numerical modelling has to be used[32]. Another interesting area would the study models with rough

[32]It has to be said that for all closed form expressions in finance, numerical results *must* be used, since it is impossible to derive an analytical expression for the cumulative distribution function for the normal distribution.

volatility. We may consider a simple case. We define

$$Y(t) = y_0 + \int_0^t (t-s)^{\alpha-1} Y(s) dW(s), \tag{141}$$

where the solution $t \mapsto Y(t) \in C^\alpha$, where C^α is a Hölder space, in the sense that $x \in C^\alpha$ if and only if

$$\|x\|_\alpha = \sup_{s,t\in[0,T]} \frac{|x(t)-x(s)|}{|t-s|^\alpha} < \infty,$$

where $\|\cdot\|_\alpha$ is a semi-norm.

We may then define the volatility v as

$$v(t) = Y(t) \otimes Y(t).$$

We note that we have to use methods from *Rough path theory* to solve equation (141), see Friz & Hairer [41] for an introduction. The idea is the same as with the stochastic volatility. That is, proving pricing theorems for contingent claims in infinite dimensional spaces.

13.4 Infinite dimensional risk measures and backward stochastic differential equations

It is known that it is possible to obtain a large set of risk measures by using an operator given by the solution of a *backward stochastic differential equation* (abbreviated as BSDE), known as the g-expectation, where g is the driver of the BSDE. BSDEs are an important class of equations in mathematical finance, most often used to price derivatives. It can be explained by considering a plain European option. In the Black-Scholes-Merton framework, in order to price such an option, we need to find a self-financing strategy that replicates the payoff of the option. Then, we know the dynamics of the strategy and its final condition. And we see that this corresponds to a BSDE. We refer to Karoui, Peng & Quenez [32] for a further discussion on this topic.

Under some given assumptions, mainly that g is Lipschitz and in L^2 and starts in 0, the BSDE

$$\begin{aligned} -dY(t) &= g\left(t, Y(t), Z(t)\right) dt - Z(t) dW(t) \\ Y(T) &= \xi, \end{aligned} \tag{142}$$

has a unique square integrable adapted solution for each $\xi \in L^2(\mathcal{F}_t)$. By a solution, we mean a pair $(Y(t), Z(t))$ that solved (142) for $0 \leq t \leq T$. Moreover, if the driver g is convex and $Y(X) = -X$, then it is possible to show that the solution $Y(t)$ of (142) is a convex risk measure, see for example **Theorem 3.2** in Barrieu & El Karoui [6] and the references therein.

This theory could be extended to infinite dimensional spaces, and in particular (if possible) to the Filipović space. One could also introduce topics from stochastic optimal control, since it is shown in Karoui, Peng & Quenez [32] that the duality between the hedging problem and the pricing problem corresponds to a general duality between BSDEs and control problems.

14 Appendices

14.1 Appendix 1 - Frequently used notation

$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sets of natural numbers, integers, rational numbers, real numbers and complex numbers

$\mathbb{R}_+$ the set of all non-negative real numbers

$\mathbb{R} - \{0\}$ the set of all real numbers sans zero

$[a, b], (a, c)$ closed and open intervals

$\bigcap A, \bigcup A$ intersection and union

$A \times B$ Cartesian product of A and B

$\mathcal{B}(\Omega)$ Borel σ-algebra of Borel sets of Ω

$\sigma(A)$, σ-algebra generated by A

$(\Omega, \mathcal{A}, \mu)$ measure space

$(\Omega, \mathcal{F}, P)$ probability space

$\mathbb{E}[X], \mathrm{Var}[X]$ expectation and variance of X

i.i.d independent and identically distributed

$\mathrm{Cov}(X, Y), \mathrm{Corr}(X, Y)$ covariance and correlation

$N(\mu, \sigma^2)$ normal distribution with mean μ and variance σ^2

$\{F_t\}$ filtration

$\mathbb{1}(A)$ indicator function over some set A

Tr trace

$\langle \cdot, \cdot \rangle$ inner product

■ end of proof

□ end of example

a.s, a.e almost surely, almost everywhere

$\log x$ natural logarithm of x

$(x - K)^+$ maximum function. Equivalent to $\max(0, x - K)$

CS Cauchy-Schwarz

BS, B76 Black-Scholes, Black-76

$\mathbb{E}[X|\mathcal{F}_t]$ conditional expectation of X given $\mathcal{F}_t$

$\sum_{k=1}^{n} \mathit{\Sigma}_k$ sum over $\mathit{\Sigma}_k$. Notice the difference! $\mathit{\Sigma}$ will be used frequently as both a sum and variable, the only difference will be the slanting.

$\hat{C}(X, X^*) = \hat{C}(X)$ the smallest σ-algebra on X with respect to which all the functionals $x^* \in X^*$ are measurable. Sometimes called the cylindrical σ-algebra of X.

14.2 Appendix 2 - Basic results on mathematical finance

The next two results are two basic results from mathematical finance. We refer to Black & Scholes [14] or Benth [9] for proofs of the unproven results.

Lemma 14.1. *The put call parity for European call and put options is*

$$P^{put}(t) = P^{call}(t) - S(t) + Ke^{-r(T-t)}. \tag{143}$$

Lemma 14.2 (Black & Scholes formula for put options)**.** *The price of a European put option is*

$$P^{put}(t) = -S(t)\Phi(-d_1) + Ke^{-rt}\Phi(-d_2). \tag{144}$$

Lemma 14.3.

$$\phi(d_1)S(t) = \phi(d_2)Ke^{-r(T-t)} \tag{145}$$

Proof:

$$\begin{aligned}
\log(\phi(d_1)) - \log(\phi(d_2)) &= \log\left(\frac{\phi(d_1)}{\phi(d_1 - \sigma\sqrt{T-t})}\right) \\
&= \log\left(e^{-\frac{1}{2}d_1^2}e^{\frac{1}{2}(d_1-\sigma\sqrt{T-t})^2}\right) \\
&= -\frac{1}{2}d_1^2 + \frac{1}{2}\left(d_1^2 + \sigma^2(T-t) - 2\sigma\sqrt{T-t}\,d_1\right) \\
&= \frac{\sigma^2}{2}(T-t) - \sigma d_1\sqrt{T-t} \\
&= -\log\left(\frac{S(t)}{K}\right) - r(T-t)
\end{aligned}$$

Hence

$$\frac{\phi(d_1)}{\phi(d_2)} = e^{-r(T-t)}\frac{K}{S(t)},$$

which in turn implies that

$$S(t)\phi(d_1) = Ke^{-r(T-t)}\phi(d_2),$$

which is what we wanted to prove. ■

Proposition 14.4. *The Black-Scholes formula is monotone as a function of the volatility.*

Proof: Without loss of generality, we do the proof for the formula for put options, which is here denoted as $P(t)$. We need to prove that the derivative of P with respect to σ is strictly positive.

$$
\begin{aligned}
\frac{\partial}{\partial \sigma} P(t) &= \frac{\partial}{\partial \sigma}\Big(-S(t)\Phi(-d_1) + Ke^{-r(T-t)}\Phi(-d_2)\Big) \\
&= \frac{\partial}{\partial \sigma}\Big(-S(t)(1-\Phi(d_1)) + Ke^{-r(T-t)}(1-\Phi(d_2))\Big) \\
&= S(t)\phi(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{-r(T-t)}\phi(d_2)\frac{\partial d_2}{\partial \sigma} \\
&= S(t)\phi(d_1)\frac{\partial d_1}{\partial \sigma} - Ke^{(}-r(T-t))\phi(d_2)(\frac{\partial d_1}{\partial \sigma} - \sqrt{T-t}) \\
&= \frac{\partial d_1}{\partial \sigma}\left(\underbrace{S(t)\phi(d_1) - Ke^{-r(T-t)}\phi(d_2)}_{=0 \text{ by } (145)}\right) + Ke^{-r(T-t)}\phi(d_2)\sqrt{T-t} \\
&= Ke^{-r(T-t)}\phi(d_2)\sqrt{T-t} > 0.,
\end{aligned}
$$

which proves the claim.

■

Proposition 14.5. *Let $B(t)$ be a Brownian motion. It then holds true that*

$$\mathbb{E}[B(t)B(s)] = \min(t,s),$$

from which it follows that $Cov(B(t), B(s)) = \min(t,s)$.

Proof: Without loss of generality, we assume that $t \geq s$.

$$
\begin{aligned}
\mathbb{E}[B_t B_s] &= \mathbb{E}[(B_t - B_s + B_s)B_s] \\
&= \mathbb{E}[(B_t - B_s)B_s + B_s^2] \\
&= s,
\end{aligned}
$$

since $(B_t - B_s)$ is independent of B_s by the independent increment property of Wiener processes.

The same is true if we assume $s \geq t$, from which the result follows. The result about covariance follows directly from the definition of the covariance

$$Cov(B_t, B_s) = \mathbb{E}[B_t B_s] - \mathbb{E}[B_t]\,\mathbb{E}[B_s].$$

■

14.3 Appendix 3 - Background on normed spaces and topological spaces

Definition 14.6 (Topology)**.** Let Ω be a set. A *topology* τ on Ω is a collection of subsets such that

1. $\emptyset$ and Ω are in τ.
2. Countable unions of any subcollection of τ is in τ.
3. Finite intersections of elements in τ is in τ.

The pair (Ω, τ) is called a *topological space*.

Definition 14.7 (Basis for a topology)**.** Let Ω be a set. A *basis* for a topology on Ω is a collection $\mathcal{B}$ of subsets of Ω such that

1. For each $x \in \Omega$, there exists at least one basis element $B \in \mathcal{B}$ containing x.
2. If $x \in B_1 \cap B_2$, then there exists a basis element $B_3 \in \mathcal{B}$ containing x such that $B_3 \subset B_1 \cap B_2$.

Definition 14.8 (Metric)**.** Let Ω be a set. A *metric* on Ω is a function $d : \Omega \times \Omega \to [0, \infty)$ such that for $x, y, z \in \Omega$

1. $d(x, y) \geq 0$.
2. $d(x, y) = d(y, x)$.
3. $d(x, y) \leq d(x, z) + d(z, y)$.
4. $d(x, y) = 0$ if and only if $x = y$.

Definition 14.9 (Norm)**.** Let X be a linear space over $\mathbb{C}$ or $\mathbb{R}$. A *norm* on X is a function $\|\cdot\| : X \to [0, \infty)$ such that

1. $\|x + y\| \leq \|x\| + \|y\|$.
2. $\|\lambda x\| = |\lambda|\|x\|$.
3. $\|x\| = 0$ if and only if $x = 0$.

All these definitions (and many more) can be found in [45] and [60].

The next definition is from McDonald and Weiss [55], and gives an alternative definition of a locally convex topological vector space.

Definition 14.10 (Locally Convex Topological Linear Space)**.** A topological linear space Ω is said to be locally convex if there exists a collection $\mathcal{W}$ of convex open subsets, each containing 0 such that

1. $W_1, W_2 \in \mathcal{W} \implies W_3 \subset W_1 \cap W_2$ for some $W_3 \in \mathcal{W}$.
2. $W \in \mathcal{W}$ and $x \in W \implies$ there exists a $W_1 \in \mathcal{W}$ such that $x + W_1 \subset W$.
3. $W \in \mathcal{W} \implies$ there exists a $W_1 \in \mathcal{W}$ such that $W_1 + W_1 \subset W$.
4. $W \in \mathcal{W} \implies$ there exists a $W_1 \in \mathcal{W}$ and an $\epsilon > 0$ such that $\alpha W_1 \subset W$ whenever $|\alpha| < \epsilon$.
5. $\{x + W : x \in \Omega, W \in \mathcal{W}\}$ is a basis for the open sets of Ω.

14.4 Appendix 4 - Table of proofs

This appendix is a list of my contributions.

My result, statement and proof	**Others result, but my proof**	**Modified versions of others proofs**
Lemma 5.4	Corollary 5.13	Proposition 5.5
Theorem 5.8	Proposition 8.4	Proposition 5.6
Theorem 5.9	Lemma 8.6	Theorem 8.36
Theorem5.10	Theorem 8.9	Theorem 8.37
Theorem5.11	Proposition 8.17	Lemma 14.3[33]
Lemma 8.3	Lemma 8.18	Proposition 14.4[34]
Proposition 8.10	Proposition 8.21	
Corollary 8.11	Lemma 8.35	
Corollary 9.10	Lemma 9.1	
Corollary 9.11	Theorem 9.4[35]	
Proposition 9.12	Lemma 9.5	
Theorem 9.13	Lemma 9.6	
Proposition 9.18	Lemma 9.8	
Corollary 9.19	Theorem 9.9	
Theorem 10.9	Lemma 9.15	
Proposition 11.4	Theorem 9.16[36]	
Theorem 12.17	Corollary 9.17	
	Theorem 10.6	
	Theorem 10.7	
	Theorem 11.13	

[36]Given as an exercise at UiO. Result assumed to be folklore.
[36]Given as an exercise at UiO. Assumed to be folklore as well.
[36]The proof in the cited article had some inconsistencies, I have therefore made a new proof.
[36]Somewhate modified statement.

14.5 Appendix 5 - List of figures

References

[1] Applebaum, D. "Infinite Dimensional Ornstein-Uhlenbeck Processes Driven by Lévy Processes". In: *Probab. Surveys* 12 (2015), pp. 33–54. DOI: 10.1214/14-PS249.

[2] Artzner, P.; Delbaen, F.; Eber, J. M., and Heath, D. "Coherent Measures of Risk". In: *Mathematical Finance* 9.3 (1999). DOI: 10.1111/1467-9965.00068.

[3] Audet, N.; Heiskanen, P.; Keppo, J., and Vehviläinen, L. "12 Modeling Electricity Forward Curve Dynamics in the Nordic Market". In: 2004.

[4] Baker, C. R. and McKeague, I. "Compact Covariance Operators". In: *Proceedings of the American Mathematical Society* 83.3 (1981), pp. 590–590. ISSN: 0002-9939,1088-6826. DOI: 10.1090/S0002-9939-1981-0627699-7.

[5] Bannör, K. F. and Scherer, M. "Capturing Parameter Risk with Convex Risk Measures". In: *European Actuarial Journal* 3 (2013), pp. 97–132.

[6] Barrieu, P. and Karoui, N. E. "Optimal Derivatives Design under Dynamic Risk Measures". In: *Contemporary Mathematics* (2004).

[7] Barth, A. and Benth, F. E. "The Forward Dynamics in Energy Markets – Infinite-Dimensional Modelling and Simulation". In: *Stochastics An International Journal of Probability and Stochastic Processes* 86.6 (2014), pp. 932–966. ISSN: 1744-2508,1744-2516. DOI: 10.1080/17442508.2014.895359.

[8] Beckers, S. "Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability". In: *Journal of Banking & Finance* 5.3 (1981). ISSN: 0378-4266. DOI: 10.1016/0378-4266(81)90032-7.

[9] Benth, F. E. *Option Theory with Stochastic Analysis: An Introduction To Mathematical Finance (Universitext)*. Springer, 2004. ISBN: 3-540-40502-X.

[10] Benth, F. E.; Benth, J., and Koekebakker, S. *Stochastic Modeling Of Electricity And Related Markets (Advanced Series on Statistical Science and Applied Probability) (Volume 11)*. World Scientific, 2008. ISBN: 981-281-230-X.

[11] Benth, F. E. and Krühner, P. "Derivatives Pricing in Energy Markets: An Infinite-Dimensional Approach". In: *SIAM Journal on Financial Mathematics* 6.1 (Jan. 2015). DOI: 10.1137/15100268X.

[12] Benth, F. E. and Krühner, P. "Representation of Infinite-Dimensional Forward Price Models in Commodity Markets". In: *Communications in Mathematics and Statistics* 2.1 (Mar. 2014). DOI: 10.1007/s40304-014-0030-1.

[13] Black, F. "The Pricing of Commodity Contracts". In: *Journal of Financial Economics* 3.1-2 (1976), pp. 167–179.

[14] Black, F. and Scholes, M. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (May 1973). DOI: 10.1086/260062.

[15] Bosq, D. *Linear Processes in Function Spaces: Theory and Applications (Lecture Notes in Statistics)*. Springer, 2000. ISBN: 0-387-95052-4.

[16] Bowley, A. L. "The Standard Deviation of the Correlation Coefficient". In: *Journal of the American Statistical Association* 23.161 (1928), pp. 31–34. ISSN: 0162-1459,1537-274X. DOI: 10.2307/2277400.

[17] Carmona, R. and Durrleman, V. "Pricing and Hedging Spread Options". In: *SIAM Review* 45.4 (Dec. 2003). DOI: 10.2307/25054458.

[18] Carmona, R. and Durrleman, V. "Pricing and Hedging Spread Options in a Log-Normal Model". In: (Mar. 16, 2003).

[19] Carmona, R. and Durrleman, V. *Pricing and Hedging Spread Options, Erronous Version*. URL: https://carmona.princeton.edu/download/fe/sirev.pdf.

[20] Carmona, R. and Tehranchi, M. R. *Interest Rate Models: An Infinite Dimensional Stochastic Analysis Perspective (Springer Finance)*. Springer, 2006. ISBN: 3-540-27065-5.

[21] Chance, D. M. *Rational Rules and Boundary Conditions for Option Pricing*. July 25, 2008.

[22] Chobanjan, S. A. and Tarieladze, V. I. "On the Compactness of Covariance Operators". In: *Soobšč. Akad Nauk Gruzin* 70 (1973), pp. 273–276.

[23] Chung, K. L. *Introduction to Stochastic Integration (Probability and Its Applications)*. Birkhäuser, Jan. 1990. ISBN: 0-8176-3386-3.

[24] Cizeau, P.; Stanley, H. E.; Liu, Y.; Meyer, M.; Peng, C. K., and Eugene, H. "Volatility Distribution in the S&P500 Stock Index". In: *Physica A: Statistical Mechanics and its Applications* 245.3-4 (1997). DOI: 10.1016/s0378-4371(97)00417-2.

[25] Cont, R. "Model Uncertainty and Its Impact on the Pricing of Derivative Instruments". In: *Mathematical Finance* 16.3 (2006). DOI: 10.1111/j.1467-9965.2006.00281.x.

[26] Dahl, K. R. *Convexity, Optimization and Risk Measures*. 2019.

[27] Delbaen, F. "Coherent Risk Measures". In: *Blätter der DGVFM* 24.4 (2000), pp. 733–739. ISSN: 1864-0281,1864-0303. DOI: 10.1007/bf02809088. URL: http://doi.org/10.1007/bf02809088.

[28] Dellacherie, C. and Meyer, P.-A. *Probabilities and Potential (Mathematics Studies, 29) (English and French Edition)*. Elsevier Science Ltd, 1979. ISBN: 0-7204-0701-X.

[29] Derman, E. *Model Risk*. 1996.

[30] Devore, J. L. and Berk, K. N. *Modern Mathematical Statistics with Applications (Springer Texts in Statistics)*. Springer, 2011. ISBN: 1-4614-0390-1.

[31] Doege, H.-J. L. J. "Convex Risk Measures for Portfolio Optimization and Concepts of Flexibility". In: *Mathematical Programming* 104.2-3 (2005), pp. 541–559. ISSN: 0025-5610,1436-4646. DOI: 10.1007/s10107-005-0628-x.

[32] El Karoui, N.; Peng, S., and Quenez, M. C. "Backward Stochastic Differential Equations in Finance". In: *Mathematical Finance* 7.1 (1997), pp. 1–71. ISSN: 0960-1627,1467-9965. DOI: 10.1111/1467-9965.00022.

[33] Filipovic, D. *Term-Structure Models: A Graduate Course (Springer Finance)*. Springer, Sept. 2009. ISBN: 3-540-09726-0.

[34] Filipović, D. "[Lecture Notes in Mathematics] Consistency Problems for Heath-Jarrow-Morton Interest Rate Models Volume 1760 || References". In: vol. 10.1007/b76888. Chapter 9. 2001. ISBN: 978-3-540-44548-7. DOI: 10.1007/978-3-540-44548-7_9.

[35] Filipović, D. and Svindland, G. "The Canonical Model for Law Invariant Convex Risk Measures Is L1". In: *Mathematical Finance* 22.3 (2012). ISSN: 0960-1627,1467-9965. DOI: 10.1111/j.1467-9965.2012.00534.x.

[36] Fisher, R. A. "Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population". In: *Biometrika* 10.4 (1915), pp. 507–521. ISSN: 0006-3444,1464-3510. DOI: 10.2307/2331838.

[37] Fisher, R. A. "On the'probable Error'of a Coefficient of Correlation Deduced from a Small Sample". In: (1921).

[38] Flyvbjerg, H. and Petersen, H. G. "Error Estimates on Averages of Correlated Data". In: *The Journal of Chemical Physics* 91.1 (1989), pp. 461–466. DOI: 10.1063/1.457480. eprint: https://doi.org/10.1063/1.457480.

[39] Folland, G. B. *Real Analysis: Modern Techniques and Their Applications*. Wiley, May 2007. ISBN: 0-471-31716-0.

[40] Föllmer, H. and Schied, A. "Convex Measures of Risk and Trading Constraints". In: *Finance and Stochastics* 6.4 (Oct. 2002). DOI: 10.1007/s007800200072.

[41] Friz, P. K. and Hairer, M. *A Course on Rough Paths: With an Introduction to Regularity Structures (Universitext)*. Springer, 2014.

[42] Heath, D.; Jarrow, R., and Morton, A. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation". In: *Econometrica* 60.1 (Jan. 1992), pp. 77–105.

[43] Hotelling, H. "New Light on the Correlation Coefficient and Its Transforms". In: *Journal of the Royal Statistical Society. Series B (Methodological)* 15.2 (1953), pp. 193–232. ISSN: 00359246. JSTOR: 2983768.

[44] Hull, J. C. *Options, Futures, and Other Derivatives, Global Edition (English and Spanish Edition)*. Pearson, 2017. ISBN: 978-1-292-21289-0.

[45] Jameson, G. J. O. *Topology and Normed Spaces*. John Wiley & Sons, Jan. 1974. ISBN: 0-470-43924-6.

[46] Jonsson, M. "Standard Error Estimation by an Automated Blocking Method". In: *Phys. Rev. E* 98.4 (Oct. 2018), p. 043304. DOI: 10.1103/PhysRevE.98.043304.

[47] Kaina, M. and Rüschendorf, L. "On Convex Risk Measures on Lp-Spaces". In: *Mathematical Methods of Operations Research* 69.3 (July 1, 2009), pp. 475–495. ISSN: 1432-5217. DOI: 10.1007/s00186-008-0248-3.

[48] Kirk, E. "Correlation in the Energy Markets, in Managing Energy Price Risk". In: *Risk Publications and Enron* (1995).

[49] Lax, P. D. *Functional Analysis*. Wiley-Interscience, Apr. 2002. ISBN: 0-471-55604-1.

[50] Li, D. X. "On Default Correlation: A Copula Function Approach". In: *SSRN Electronic Journal* (1999). ISSN: 1556-5068. DOI: 10.2139/ssrn.187289.

[51] Lindström, E. "Implications of Parameter Uncertainty on Option Prices". In: *ADS* 2010 (Mar. 2010). DOI: 10.1155/2010/598103.

[52] Lindstrøm, T. L. *Spaces: An Introduction to Real Analysis (Pure and Applied Undergraduate Texts)*. American Mathematical Society, 2017. ISBN: 1-4704-4062-8.

[53] Margrabe, W. "The Value of an Option to Exchange One Asset for Another". In: *Journal of Finance* 33.1 (1978), pp. 177–86.

[54] Mayhew, S. "Implied Volatility". In: *Financial Analysts Journal* 51.4 (1995), pp. 8–20. ISSN: 0015-198X. DOI: 10.2307/4479853.

[55] McDonald, J. N. *A Course in Real Analysis*. Academic Press, Jan. 2012. ISBN: 0-12-387774-1.

[56] Meiss, J. D. *Differential Dynamical Systems (Monographs on Mathematical Modeling and Computation)*. Society for Industrial and Applied Mathematics, 2007. ISBN: 0-89871-635-7.

[57] Merton, R. C. "An Intertemporal Capital Asset Pricing Model". In: *Econometrica* 41.5 (1973), pp. 867–887. ISSN: 0012-9682,1468-0262. DOI: 10.2307/1913811.

[58] Merton, R. C. "The Relationship between Put and Call Option Prices: Comment". In: *The Journal of Finance* 28.1 (1973), pp. 183–184. ISSN: 0022-1082,1540-6261. DOI: 10.2307/2978180.

[59] Merton, R. C. "Theory of Rational Option Pricing". In: *The Bell Journal of Economics and Management Science* 4.1 (1973), pp. 141–183. ISSN: 0005-8556. DOI: 10.2307/3003143.

[60] Munkres, J. *Topology (2nd Edition)*. Pearson, Jan. 2000. ISBN: 0-13-181629-2.

[61] Musiela, M. "Stochastic PDEs and Term Structure Models". In: *Journées Internationales de Finance* (1993).

[62] Oksendal, B. *Stochastic Differential Equations*. Springer Verlag, Aug. 2019. ISBN: 3-540-04758-1.

[63] Olkin, I. and Pratt, J. W. "Unbiased Estimation of Certain Correlation Coefficients". In: *The Annals of Mathematical Statistics* 29.1 (1958), pp. 201–211. ISSN: 0003-4851. DOI: 10.2307/2237306.

[64] Peszat, S. and Zabczyk, J. *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach (Encyclopedia of Mathematics and Its Applications)*. Cambridge University Press, 2007. ISBN: 0-521-87989-2.

[65] Prato, G. D. and Zabczyk, J. *Stochastic Equations in Infinite Dimensions (Encyclopedia of Mathematics and Its Applications)*. Cambridge University Press, 2008. ISBN: 0-521-05980-1.

[66] Protter, P. E. *Stochastic Integration and Differential Equations*. Springer, 2005. ISBN: 978-3-540-00313-7.

[67] Reisinger, C.; Whitley, A., and Gupta, A. “Model Uncertainty and Its Impact on Derivative Pricing”. In: 2010.

[68] Rockafellar, R. T. “Coherent Approaches to Risk in Optimization under Uncertainty”. In: *OR Tools and Applications: Glimpses of Future Technologies*. 2007. 2008. Chap. Chapter 3, pp. 38–61. DOI: 10.1287/educ.1073.0032.

[69] Ruszczyński Andrzej; Shapiro, A. “Corrigendum to: “Optimization of Convex Risk Functions,” Mathematics of Operations Research 31 (2006) 433–452”. In: *Mathematics of Operations Research* 32.2 (2007), pp. 496–496. ISSN: 0364-765X,1526-5471. DOI: 10.1287/moor.1070.0265.

[70] Ruszczyński, A. and Shapiro, A. “Optimization of Convex Risk Functions”. In: *Mathematics of Operations Research* 31.3 (2006), pp. 433–452. ISSN: 0364-765X,1526-5471. DOI: 10.1287/moor.1050.0186.

[71] Rynne, B. and Youngson, M. A. *Linear Functional Analysis (Springer Undergraduate Mathematics Series)*. Springer, Dec. 2007. ISBN: 1-84800-004-9.

[72] Schwartz, E. S. “The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging”. In: *The Journal of Finance* 52.3 (1997), pp. 923–973. ISSN: 0022-1082,1540-6261. DOI: 10.1111/j.1540-6261.1997.tb02721.x.

[73] Vakhania, N. N. and Tarieladze, V. I. “Covariance Operators of Probability Measures in Locally Convex Spaces”. In: *Theory of Probability & Its Applications* 23.1 (Dec. 1978). DOI: 10.1137/1123001.

[74] Vakhania, N.; Tarieladze, V., and Chobanyan, S. *Probability Distributions on Banach Spaces (Mathematics and Its Applications)*. Springer, 1987. ISBN: 90-277-2496-2.

[75] Wiersema, U. F. *Brownian Motion Calculus*. Wiley, 2008. ISBN: 0-470-02170-5.

[76] Zimmermann, D. W.; Zumbo, B., and Williams, R. H. “Bias in Estimation and Hypothesis Testing of Correlation”. In: *Psicológica* 24 (Jan. 2003).

Index